



US009328397B2

(12) **United States Patent**
Drewniak et al.

(10) **Patent No.:** US 9,328,397 B2
(45) **Date of Patent:** May 3, 2016

(54) **REMOVAL OF ARSENIC USING A DISSIMILATORY ARSENIC REDUCTASE**

(71) Applicant: **UNIWERSYTET WARSZAWSKI**, Warszawa (PL)

(72) Inventors: **Lukasz Drewniak**, Skarzysko-Kamienna (PL); **Aleksandra Sklodowska**, Warszawa (PL); **Monika Radlinska**, Warszawa (PL); **Robert Stasiuk**, Sarnaki (PL)

(73) Assignee: **UNIWERSYTET WARSZAWSKI**, Warsaw (PL)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.

(21) Appl. No.: **14/678,143**

(22) Filed: **Apr. 3, 2015**

(65) **Prior Publication Data**

US 2015/0267276 A1 Sep. 24, 2015

Related U.S. Application Data

(63) Continuation-in-part of application No. PCT/IB2013/059773, filed on Oct. 30, 2013.

(30) **Foreign Application Priority Data**

Jun. 19, 2013 (PL) 404376

(51) **Int. Cl.**

C12N 15/63 (2006.01)
C22B 3/18 (2006.01)
C12N 9/02 (2006.01)
C07K 14/195 (2006.01)
C12N 9/14 (2006.01)
C02F 3/34 (2006.01)
C12P 3/00 (2006.01)
C12R 1/01 (2006.01)
C22B 1/11 (2006.01)
C02F 3/28 (2006.01)
C02F 101/10 (2006.01)

(52) **U.S. Cl.**

CPC ... *C22B 3/18* (2013.01); *C02F 3/34* (2013.01);
C02F 3/342 (2013.01); *C07K 14/195* (2013.01); *C12N 9/0004* (2013.01); *C12N 9/14* (2013.01); *C12P 3/00* (2013.01); *C12R 1/01* (2013.01); *C12Y 120/99001* (2013.01); *C12Y 306/03016* (2013.01); *C22B 1/11* (2013.01);

C02F 3/28 (2013.01); *C02F 2101/103* (2013.01); *C12Y 120/04001* (2013.01); *Y02P 20/52* (2015.11)

(58) **Field of Classification Search**

None
See application file for complete search history.

(56) **References Cited**

PUBLICATIONS

Ma et al. Rejection of arsenic minerals in sulfide flotation—A literature review. International Journal of Mineral Processing vol. 93, Issue 2, Oct. 1, 2009, pp. 89-94.*

L. Drewniak, et al., Arsenic Release From Gold Mine Rocks Mediated by the Activity of Indigenous Bacteria, Hydrometallurgy (2010) vol. 104, p. 437-442.

Lukasz Drewniak, et al., The Contribution of Microbial Mats to the Arsenic Geochemistry of an Ancient Gold Mine, Environmental Pollution (2012) vol. 162, p. 190-201.

X. Ma, et al., Rejection of Arsenic Minerals in Sulfide Flotation—A Literature Review, Int. J. Miner. Process., (2009) vol. 93, p. 89-94.

D. Malasarn, et al., arrA Is a Reliable Marker for As(V) Respiration, Science, Oct. 15, 2004, vol. 306, p. 455.

Joanne M. Santini, et al., A New Chemolithoautotrophic Arsenic-Oxidizing Bacterium Isolated From a Gold Mine: Phylogenetic, Physiological, and Preliminary Biochemical Studies, Applied and Environmental Microbiology (2000) vol. 66, No. 1, p. 92-97.

Database EMBL: Accession No. HQ316511: May 1, 2011, Uncultured Bacterium clone ZSARR21 Arsenate Respiratory Reductase-Like (arrA) Gene, partial sequence.

Database EMBL: Accession No. AY660886: Oct. 2, 2004, *Shewanella* sp. HAR-4 Arsenate Respiratory Reductase (arrA) Gene, partial cds.

International Search Report and Written Opinion for PCT/IB2013/059773 dated Oct. 28, 2914.

* cited by examiner

Primary Examiner — Michele K Joike

(74) **Attorney, Agent, or Firm** — Vedder Price P.C.; Thomas J. Kowalski; Deborah L. Lu

(57) **ABSTRACT**

The object of the invention is the plasmid pSheB, particularly a plasmid which may comprise a fragment of pSheB including the arr module and functional derivatives thereof, and strains containing such a plasmid, preferably the *Shewanella* sp. strain, deposited as KKP 2045p, which are capable of removing arsenic by dissolution of minerals and reduction of arsenates to arsenites. The object of the invention is also the method and the use of such bacterial strains or compositions which may comprise them, for the selective removal of arsenic from mineral resources, raw materials industry waste or soil.

21 Claims, 5 Drawing Sheets

Fig. 1

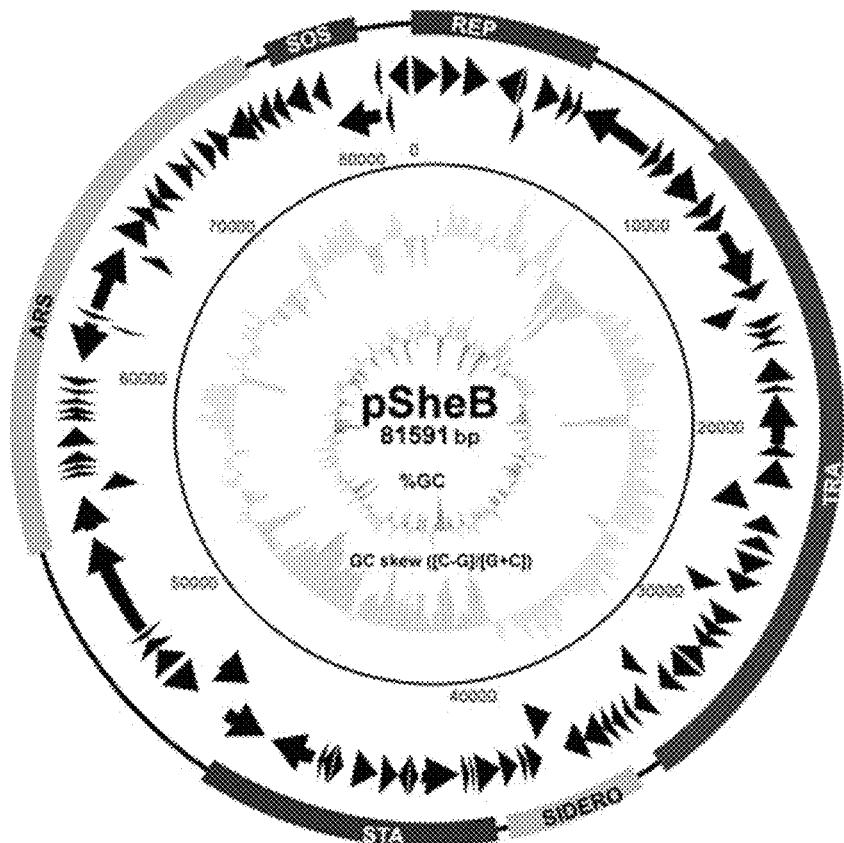


Fig. 2

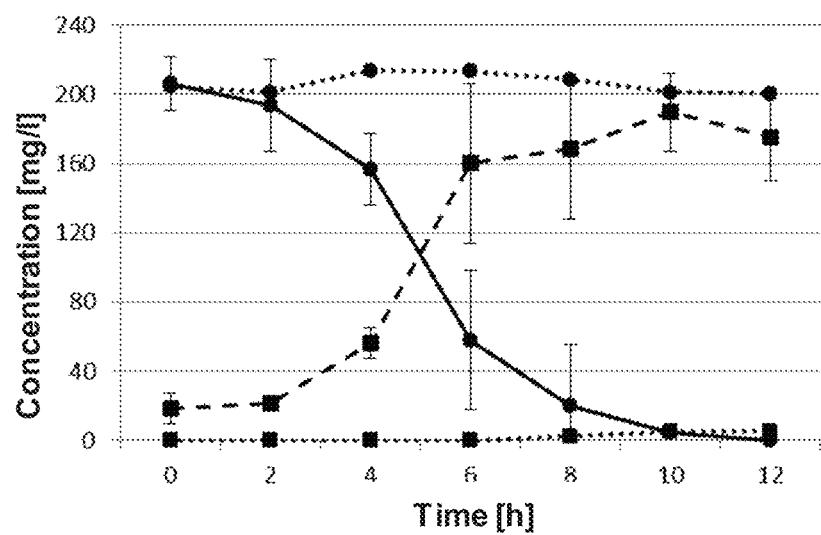


Fig. 3A-B

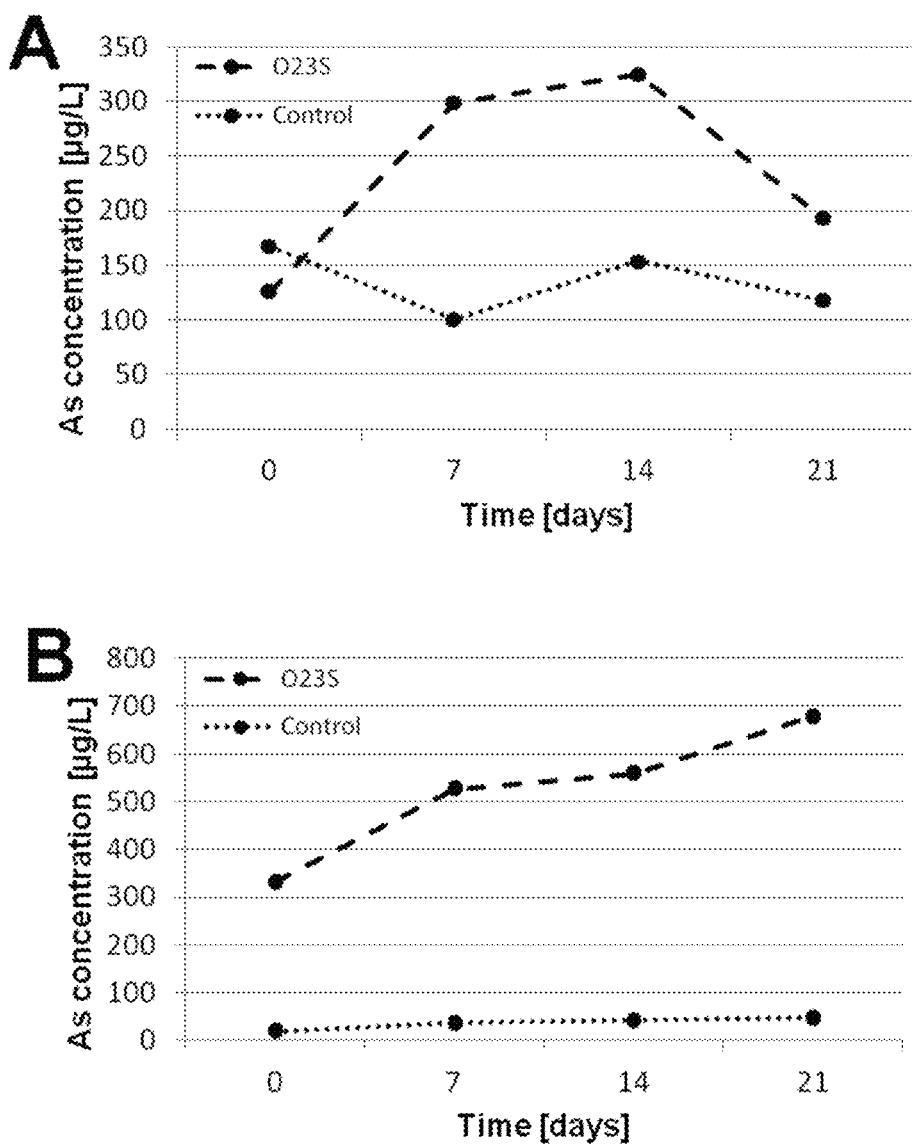


Fig. 4

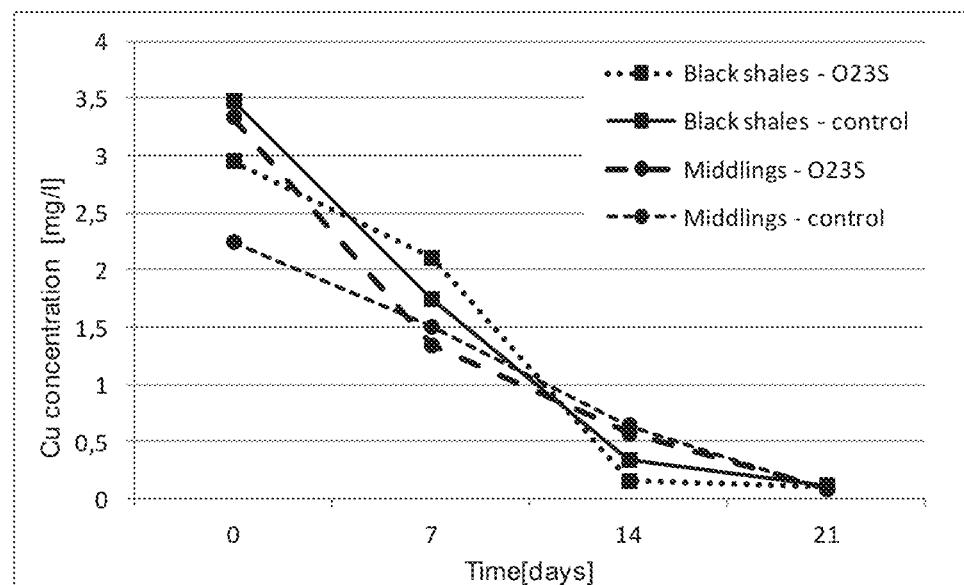


Fig. 5

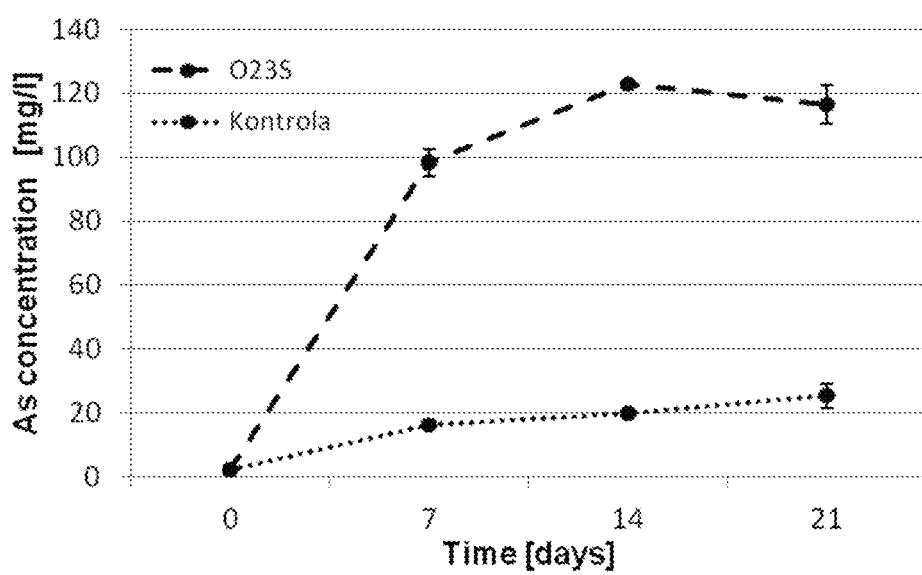


Fig. 6

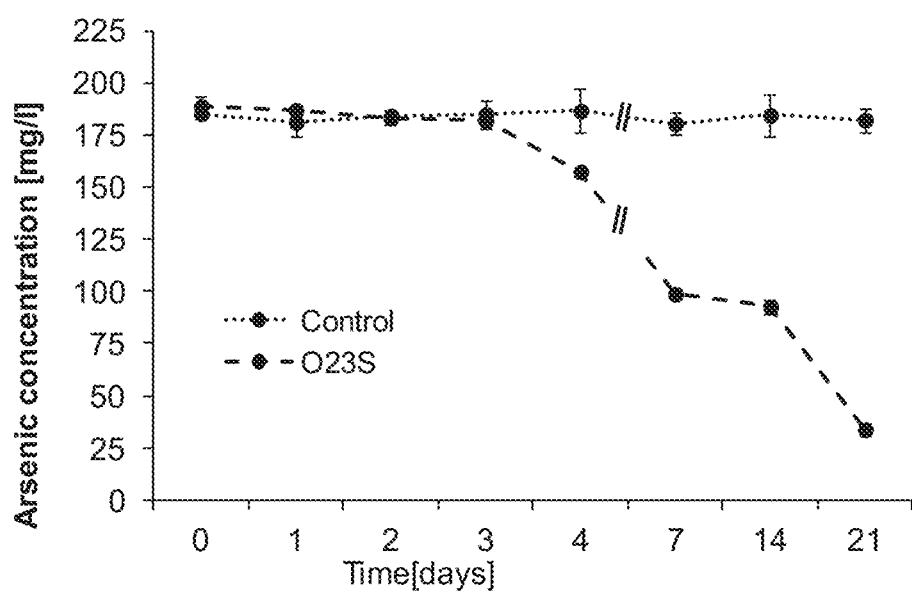
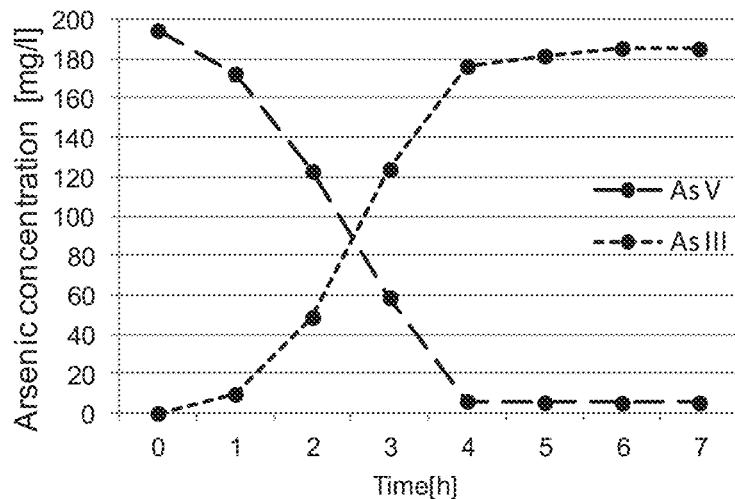
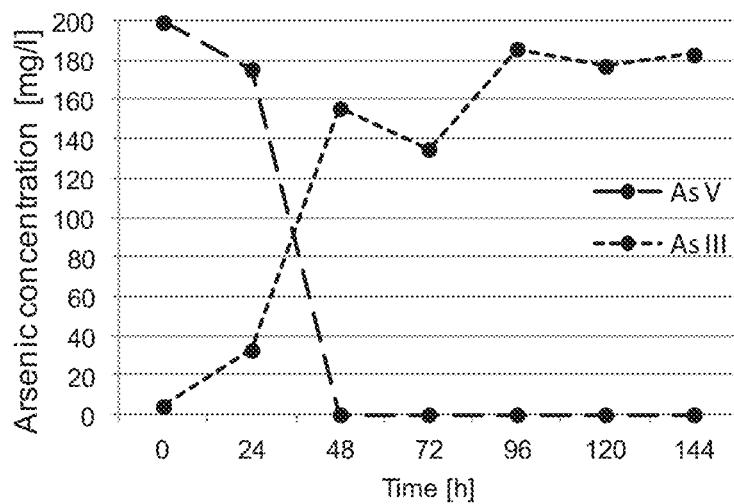


Fig. 7A-B

A**B**

1

**REMOVAL OF ARSENIC USING A
DISSIMILATORY ARSENIC REDUCTASE**

**RELATED APPLICATIONS AND
INCORPORATION BY REFERENCE**

This application is a continuation-in-part application of international patent application Ser. No. PCT/IB2013/059773 filed 30 Oct. 2013, which published as PCT Publication No. WO 2014/203046 on 24 Dec. 2014, which claims benefit of Polish patent application Serial No. P.404376 filed 19 Jun. 2013.

The foregoing applications, and all documents cited therein or during their prosecution (“appln cited documents”) and all documents cited or referenced in the appln cited documents, and all documents cited or referenced herein (“herein cited documents”), and all documents cited or referenced in herein cited documents, together with any manufacturer’s instructions, descriptions, product specifications, and product sheets for any products mentioned herein or in any document incorporated by reference herein, are hereby incorporated herein by reference, and may be employed in the practice of the invention. More specifically, all referenced documents are incorporated by reference to the same extent as if each individual document was specifically and individually indicated to be incorporated by reference.

SEQUENCE LISTING

The instant application contains a Sequence Listing which has been submitted electronically in ASCII format and is hereby incorporated by reference in its entirety. Said ASCII copy, created on May 19, 2015, is named 46538.00.2005_SL.txt and is 107,087 bytes in size.

FIELD OF THE INVENTION

The object of the invention is the plasmid pSheB, particularly a plasmid which may comprise a fragment of pSheB which may comprise the arr module or functional derivatives thereof, and strains which may comprise such a plasmid, particularly *Shewanella* sp. O23S strain, which are capable of removing arsenic from mineral resources, raw materials industry waste and soils, in particular, post-mining (e.g. soils) under anaerobic conditions, by dissolution of minerals and reduction of arsenates to arsenites. The object of the invention is also the method and the use of such bacterial strains or a composition which may comprise them, for the selective removal of arsenic from mineral resources, raw materials industry waste or the soil.

BACKGROUND OF THE INVENTION

Arsenic is an element that very often co-occurs in copper minerals and constitutes their specific impurity. In pyrometallurgical processes of roasting and smelting of copper concentrates, volatile arsenic compounds are released into the atmosphere, which, due to the toxicity of these compounds, constitutes a major threat to the environment. During the smelting process, most of the arsenic is removed as a volatile compound As_4O_6 at concentrations up to 0.5 mg/l, while only 0.04%-0.06% is removed in a solid, stable form with slags [Piret, 1999]. Apart from the volatile arsenic compounds, high concentrations of arsenic are also found in dusts. For this reason, it is very important from both an economic and an

2

environmental point of view, to develop an effective method of controlled removal of arsenic from copper deposits and the products of their processing.

In order for the removal of arsenic from copper minerals to bring the expected economic and environmental benefits, this process must be conducted at the early stages of the copper deposits processing, such as flotation. The traditional copper flotation systems are insufficient and inadequate for separation and division of sulfide minerals containing arsenic (e.g. 10 enargite Cu_3AsS_4 or tennantite $(\text{Cu},\text{Fe})_{12}\text{As}_4\text{S}_{13}$) from copper sulfides not containing arsenic, present in the ores. An aid to the conventional flotation methods are the methods proposed in recent years. One of the methods relates to the selective oxidation of sulfides, based on the electrochemical properties of the separated compounds [Fornasiero et al., 15 2001]. Another method of selective flotation utilizes the differences in the flotation pulp potentials [Guo and Yen, 2005]. Using the separation method based on the difference in the pulp’s potentials, minerals containing arsenic, e.g. enargite 20 (Cu_3AsS_4), can be separated from copper sulfides not containing arsenic. As a result of these processes, two fractions of concentrates are produced: (i) with a low arsenic content, and (ii) with a high arsenic content. The former can be used in pyrometallurgical processes, whereas the latter fraction of 25 concentrates, containing copper minerals contaminated with arsenic, still requires adequate treatment [Senior et al., 2006].

One of the ways that can help to solve the problem of removal of arsenic from copper minerals, is the application of biohydrometallurgical methods, using microorganisms to 30 recover metals from minerals and deposits. The use of microorganisms for the extraction of copper or gold from their ores and concentrates, is a well-known process, and is often described in the literature [Xia, L. et al., 2010; Xia, L. et al., 2009; Olson et al., 2003; Rawlings and Johnson, 2007]. Most 35 of the biohydrometallurgical processes are based on the processes of oxidation of minerals, and lead to (i) an increase in the accessibility to chemical solvents (biooxidation) or (ii) their direct dissolution (bioleaching) [Rawlings and Johnson, 2007]. Unfortunately, these methods are non-specific, 40 because they are based on the oxidation of sulfur and/or iron from minerals, and are associated with the release of all the metals associated with this type of minerals. A further limitation of the traditional biohydrometallurgical methods is the use of acidophilic bacteria in leaching from neutral or slightly 45 alkaline deposits, which is often inefficient, and sometimes even impossible, due to the need for the use of considerable amounts of sulfuric acid to acidify the deposits. In the literature microorganisms are described, mainly chemolithoautotrophic, sulfur oxidizing bacteria, belonging to the genus: 50 *Thiobacillus*, *Halothiobacillus*, *Thiomonas*, and iron oxidizing bacteria, such as: *Galonella feruginea* or *Leptothrix ochracea*, *Thiothrix* and *Beggiatoa*, which can be used in bioleaching processes at neutral pH, but are very difficult to cultivate and are still poorly understood. Furthermore, 55 bioleaching with the use of these microorganisms is time-consuming and these processes are carried out as long as several months [Sklodowska and Madakowska, 2007]. A confirmation of the lack of suitable microorganisms, capable of recovering metals under neutral or slightly alkaline conditions is the current situation in the mining market. Currently, there are no known and commercially available biotechnological methods of removing precious metals, occurring in the form of sulphides, under the conditions of neutral or slightly 60 alkaline deposits.

An alternative to the oxidation processes are selective bioreduction processes, in which the selected elements, associated with the metabolic activity of microorganisms, are

released. Although there are several examples of application of the bioreduction processes, methods for removing arsenic from copper minerals using microbial reduction are unknown. Many strains of bacteria that dissimilatory reduce arsenates have been identified, but the use of most of them is limited to the transformation of soluble arsenic compounds [Newman et al., 1998] or secondary arsenic minerals, resulting from iron compounds [Zobrist et al., 2000]. Strains capable of removing arsenic from copper concentrates and flotation tailings have also been described [Mantur et al., 2011], but the dissimilatory arsenate reduction process, carried out by the aforementioned strains, is not fully balanced and part of the arsenic may be removed out of the cells in the form of volatile, toxic arsenic compounds (unpublished data). Apart from that, the strains described by Mantur et al., 2011, simultaneously release copper and arsenic from minerals, thus lowering the value of the obtained copper concentrate.

Citation or identification of any document in this application is not an admission that such document is available as prior art to the present invention.

SUMMARY OF THE INVENTION

In the light of the described state of the art, the aim of the present invention is to overcome the indicated inconveniences and to provide a plasmid which may comprise genetic information ensuring the capability of dissimilatory arsenate reduction and selective removal of arsenic, particularly from copper deposits. The aim of the invention is to provide novel bacterial strains which may comprise such a plasmid, compositions which may comprise them and uses thereof, and the methods for the selective removal of arsenic using such strains. Furthermore, it is desirable for such microorganisms, possessing the capability of arsenate reduction, not to produce volatile, and, at the same time, toxic arsenic compounds. The *Shewanella* sp. O23S strain, that has been isolated from microbial mats from a gold mine in Zloty Stok, possesses such properties [Drewniak, 2009]. The *Shewanella* sp. O23S strain has been deposited on 24 Jul. 2012 in the IAFB Collection of Industrial Microorganisms Institute of Agricultural and Food Biotechnology in Warsaw, Poland under the deposit number KKP 2045p. This strain is capable of anaerobic growth using arsenates as the final electron acceptor and can mobilize arsenic from rocks from the gold mine in Zloty Stok [Drewniak et al., 2010]. It was unexpectedly found, that these properties are ensured by the plasmid pSheB, isolated from *Shewanella* sp. O23S, the sequence of which has been shown in SEQ ID NO: 1, particularly its region which may comprise the fragment from 63978 to 72599, which encodes, among others, dissimilatory arsenate reductase, determining these properties.

Accordingly, it is an object of the invention not to encompass within the invention any previously known product, process of making the product, or method of using the product such that Applicants reserve the right and hereby disclose a disclaimer of any previously known product, process, or method. It is further noted that the invention does not intend to encompass within the scope of the invention any product, process, or making of the product or method of using the product, which does not meet the written description and enablement requirements of the USPTO (35 U.S.C. §112, first paragraph) or the EPO (Article 83 of the EPC), such that Applicants reserve the right and hereby disclose a disclaimer of any previously described product, process of making the product, or method of using the product. It may be advanta-

geous in the practice of the invention to be in compliance with Art. 53(c) EPC and Rule 28(b) and (c) EPC. Nothing herein is to be construed as a promise.

It is noted that in this disclosure and particularly in the claims and/or paragraphs, terms such as "comprises", "comprised", "comprising" and the like can have the meaning attributed to it in U.S. Patent law; e.g., they can mean "includes", "included", "including", and the like; and that terms such as "consisting essentially of" and "consists essentially of" have the meaning ascribed to them in U.S. Patent law, e.g., they allow for elements not explicitly recited, but exclude elements that are found in the prior art or that affect a basic or novel characteristic of the invention.

These and other embodiments are disclosed or are obvious from and encompassed by, the following Detailed Description.

DEPOSIT

The *Shewanella* sp. O23S strain has been deposited on 24 Jul. 2012 in the IAFB Collection of Industrial Microorganisms Institute of Agricultural and Food Biotechnology in Warsaw, Poland under the deposit number KKP 2045p.

The Deposits with IAFB Collection of industrial Microorganisms institute of Agricultural and Food Biotechnology in Warsaw, Poland, under deposit accession number KKP 2045p were made pursuant to the terms of the Budapest Treaty. Upon issuance of a patent, all restrictions upon the deposit will be removed, and the deposit is intended to meet the requirements of 37 CFR §§1.801-1.809. The deposit will be irrevocably and without restriction or condition released to the public upon the issuance of a patent. The deposit will be maintained in the depository for a period of 30 years, or 5 years after the last request, or for the effective life of the patent, whichever is longer, and will be replaced if necessary during that period.

BRIEF DESCRIPTION OF THE DRAWINGS

The following detailed description, given by way of example, but not intended to limit the invention solely to the specific embodiments described, may best be understood in conjunction with the accompanying drawings.

For a better understanding of the invention, it has been illustrated in the examples of embodiments and in the accompanying figures, in which:

FIG. 1. Shows the genetic organization of the plasmid pSheB. In the diagram, different modules of the plasmid's structure and phenotypic regions have been described: REP/STA—replication-stabilization module, TA—toxin/antitoxin module, TRA—conjugation module, SIDERO—siderophore production module, STA—additional stabilization module, ARS—arsenic metabolism module, and SOS—SOS repair system module.

FIG. 2. Shows a comparison of the capability of dissimilatory arsenate reduction by a wild-type strain (wt) *Shewanella* sp. O23S deposited as KKP2045p (harbouring the plasmid pSheB) and its derivative, deprived of the plasmid pSheB. In order to compare the abilities of the investigated strains to reduce arsenites to arsenites, anaerobic cultures were carried out in minimal R1-R2 medium containing 2.5 mM (187.5 ppm) of sodium arsenite. As(V) and As(III) content in culture fluids collected from the cultures every 24 hours is shown on the graph. ●—indicates the concentration of As(V), ■—indicates the concentration of As(III), a solid line indicates the kinetics of As(V) reduction, carried out by

the wild-type strain, while a dashed line—by the derivative, deprived of the plasmid pSheB.

FIG. 3A-B. Shows a graph illustrating the efficiency of the process of arsenic removal from mineral resources by the *Shewanella* sp. O23S strain for A) flotation tailings, B) bituminous shales. Sterile R1-R2 media, enriched with the respective mineral resources: (A) flotation tailings (“middlings”) and (B) bituminous shales (“shales”), were used as the control samples.

FIG. 4. Shows a graph illustrating the efficiency of the process of copper removal from mineral resources by the *Shewanella* sp. O23S strain for flotation tailings (“middlings”) and bituminous shales (“shales”). Sterile R1-R2 media, enriched with the respective mineral resources: flotation tailings (“middlings”) and bituminous shales (“shales”), were used as the control samples.

FIG. 5. Shows a graph illustrating the efficiency of the process of arsenic removal from the soil contaminated with arsenic by the *Shewanella* sp. O23S strain. In order to confirm the ability of the investigated strain to remove arsenic from the soil, a soil sample originating from the Zloty Potok area was used, and an anaerobic culture in minimal R1-R2 medium was carried out. Arsenic content in culture fluids collected from the cultures every 24 hours is shown on the graph. Sterile R1-R2 media with the addition of the soil were used as the controls.

FIG. 6. Shows a graph illustrating the efficiency of arsenic removal by the *Shewanella* sp. O23S strain in a medium containing 2.5 mM (187.5 ppm) sodium arsenate and 5 mM sodium thiosulfate. Sterile R1-R2 media with sodium arsenate and sodium thiosulfate were used as the controls.

FIG. 7A-B. Shows a comparison of the capability of dissimilatory arsenate reduction by the *Shewanella* sp. O23S strain in minimal medium R1-R2, having a pH 8 (A) and pH 4 (B). In order to compare the abilities of the investigated strain to dissimilatory reduce arsenates under various pH conditions, anaerobic cultures in minimal R1-R2 medium containing 2.5 mM (187.5 ppm) sodium arsenate and 5 mM sodium lactate were carried out.

DETAILED DESCRIPTION OF THE INVENTION

Bacterial strains which may comprise a plasmid including the fragment from 63978 to 72599 of SEQ ID NO: 1, plasmid pSheB or their functional derivatives, particularly the *Shewanella* sp. O23S strain (deposited as KKP 2045p), are capable of dissimilatory arsenate reduction and selective removal of arsenic, particularly from mineral resources, raw materials industry waste and from the soil, particularly preferable is that they are capable of dissimilatory arsenate reduction and selective removal of arsenic from copper deposits, preferably under neutral or slightly alkaline conditions. The bacterial strains which may comprise the plasmid which may comprise the fragment from 63978 to 72599 of SEQ ID NO: 1, pSheB or their functional derivatives, particularly the *Shewanella* sp. O23S strain (deposited as KKP 2045p) do not produce toxic, volatile arsenic compounds and stably persist in the environment.

The invention therefore relates to the isolated plasmid which may comprise the fragment of the nucleotide sequence from 63978 to 72599 of the plasmid pSheB, having the sequence shown in SEQ ID NO: 1 or its functional derivative.

The invention also relates to the plasmid pSheB shown in SEQ ID NO: 1 or its functional derivative.

The term “functional derivative of the plasmid” or “functional derivative of the sequence” may comprise plasmids/ sequences having a nucleotide sequence coding for open

reading frames, which encode products which may comprise an amino-acid or a nucleotide sequence identical or highly homologous to the sequences coded by the indicated sequences, wherein the coding sequences or other sequences of the plasmid/sequence have been modified e.g. by substitution, replacement, deletion or insertion, such that it does not essentially alter the activity of the products of these open reading frames, and enables the maintenance of functional features carried by such plasmid/sequence. The indicated sequence will therefore be the region which may comprise the fragment from 63978 to 72599 of SEQ ID NO: 1, equally preferably it will be the sequence of the plasmid pSheB shown in SEQ ID NO: 1. A highly homologous sequence means that the sequence is homologous, preferably identical in at least 70%, preferably 80%, more preferably 90%, most preferably, in at least 95%. The term “functional derivative of the plasmid” means, therefore, plasmids having a nucleotide sequence coding for open reading frames, encoding products which may comprise an amino-acid or a nucleotide sequence identical or highly homologous to the sequences coded by the fragment from 63978 to 72599 of SEQ ID NO: 1 and/or to the sequence of the plasmid pSheB shown in SEQ ID NO: 1, wherein the coding sequences or other plasmid sequences have been modified e.g. by substitution, replacement, deletion or insertion, such that it does not essentially alter the activity of the products of these open reading frames, and enables the maintenance of functional features carried by such a plasmid.

The essence of the present invention is thus based on an unexpected finding, that it is possible to use a strain containing the plasmid which may comprise the nucleotide fragment from 63978 to 72599 of SEQ ID NO: 1, pSheB shown in SEQ ID NO: 1 or their functional derivatives, particularly the *Shewanella* sp. O23S strain, deposited as KKP 2045p, for the selective removal of arsenic from mineral resources, raw materials industry: waste and soils, preferably under neutral or slightly alkaline conditions. It was unexpectedly found, that the strain which may comprise the plasmid pSheB, shown in SEQ ID NO: 1, particularly the *Shewanella* sp. O23S strain, is: capable of (i) growth in mineral media containing bituminous black shales, flotation tailings, and post-mining soils containing arsenic, (ii) selective release of arsenic from copper minerals containing arsenic, (iii) tolerating the toxic effects of heavy metals released as a result of dissolution of minerals, is (iv) lacking the ability to produce volatile arsenic compounds, is (v) lacking the ability to mobilize copper from mineral resources. The present invention, therefore, also relates to the *Shewanella* sp. O23S strain, which may comprise the plasmid pSheB, deposited under the number KKP2045p in the IAFB Collection of Industrial Microorganisms Institute of Agricultural and Food Biotechnology in Warsaw.

The invention also relates to a composition which may comprise the isolated plasmid according to the invention and/or a bacterial strain according to the invention or a combination thereof.

The invention also relates to use of a bacterial strain according to the invention, a composition according to the invention, for the selective removal of arsenic from mineral resources, raw materials industry waste or the soil.

Particularly preferred is the use of the bacterial strain, *Shewanella* sp. O23S, which may comprise the natural plasmid pSheB, carrying: (i) all the genes necessary for dissimilatory arsenate reduction, (ii) arsenite and arsenate resistance genes, and (iii) genes coding for the replication-stabilization system for selective removal of arsenic from mineral resources, raw materials industry waste and post-mining

soils. The complete sequence of the plasmid pSheB of *Shewanella* sp. O23S has been shown in SEQ ID NO: 1.

The presented solutions according to the invention enable the removal of arsenic from mineral resources, raw materials waste, and post-mining soils, preferably under neutral or slightly alkaline conditions, preferably at a pH in the range of about 6 to about 8, using a strain which may comprise the plasmid pSheB or its derivative, more preferably *Shewanella* sp. O23S, without the need for acidification of the "environment" and without the risk of releasing toxic, volatile arsenic compounds. By the invention, it is possible to selectively remove arsenic without the undesirable release of the target metals, e.g. copper or gold.

The invention therefore relates to the method for selective removal of arsenic from mineral resources, raw materials industry waste, or the soil, in which the dissimilatory arsenate reduction step is carried out using a bacterial strain according to the invention and/or a composition according to the invention. Preferably, the dissimilatory arsenate reduction step is carried out under neutral or slightly alkaline conditions. Preferably, the mineral resources are copper deposits.

The invention also relates to the method for selective removal of arsenic from various mineral resources, raw materials industry, or soils, in which the removal of arsenic is carried out by dissimilatory arsenate reduction, using a bacterial strain according to the invention, preferably which may comprise the plasmid pSheB having the sequence shown in SEQ ID NO: 1 or its functional derivative, more preferably the *Shewanella* sp. O23S strain, deposited as KKP2045p, which method may comprise the following steps:

- a) preparation of the mineral resources, wastes or soils and mixing with an appropriate culture medium enabling the cultivation of the strain,
- b) addition of an inoculum of the strain and carrying out the culture under conditions enabling its growth and conduction of dissimilatory arsenate reduction.

It is preferable when step b) is followed by step c) of selective removal of arsenic, released from the solutions obtained in step b). Such removal of arsenic from the solutions (culture fluids), will preferably be carried out using the already developed methods, e.g. by flotation, preferably by selective precipitation of arsenites with sulfides, as a consequence of which a stable, water-insoluble compound, arsenic sulfide As_2S_3 is formed [Robins, 1985].

In the preferred method, step a) is carried out by: (i) shredding and fractionation of mineral resources, wastes and soils, preferably to the fraction of 125-250 μm (having the average particle size), due to the highest performance obtainable in this range of particle sizes. Equally preferred is (ii) the preparation of an appropriate culture medium R1-R2 with additives of suitable substrates, among which are: sodium lactate as the source of carbon and energy, yeast extract as an additional source of carbon and vitamins. Tuovinen's mineral salts (Tuovinen slats (Tuovinen and Kelly, 1974)) as the source of microelements. Furthermore, in the preferred method, the medium does not contain NO_3^- and Fe^{3+} .

It is equally preferred if in step b) the culture is carried out under conditions of an appropriate, anaerobic atmosphere, which is obtained by conducting the culture with flushing of the medium with a mixture of gases $\text{N}_2:\text{CO}_2$, preferably in a ratio 4:1. The flushing of the medium with the mixture of gases $\text{N}_2:\text{CO}_2$, preferably in a ratio 4:1, can be equally preferred already in step a) of the method of dissimilatory arsenate reduction.

In the preferred method, in order to obtain the highest performance for the *Shewanella* sp. O23S strain, step b) is carried out at a temperature in the range from 15°C. to 30°C.,

preferably about 22°C. with shaking (160 rpm), for at least 21 days. It is preferred if the density of the culture at the beginning of the process is at least about 10^6 cells/ml. In the preferred method, the inoculum of the *Shewanella* sp. O23S strain is flushed several times with saline solution or is passed several times to R1-R2 medium, enriched with sodium arsenite.

Publications cited in the description, and the references given therein, are in their entirety incorporated herein as references.

The following examples are presented merely to illustrate the invention and to clarify its various aspects, but are not intended to be limitative, and should not be equated with all its scope, which is defined in the appended claims.

In the following examples, unless it was otherwise indicated, standard materials and methods described in Sambrook et al., 2001. Molecular cloning: A laboratory manual. Cold Spring Harbor Laboratory Press, New York. were used, or the manufacturers' instructions for specific materials and methods were followed.

Although the present invention and its advantages have been described in detail, it should be understood that various changes, substitutions and alterations can be made herein without departing from the spirit and scope of the invention as defined in the appended claims.

The present invention will be further illustrated in the following Examples which are given for illustration purposes only and are not intended to limit the invention in any way.

EXAMPLES

Example 1

Characteristics of the Plasmid pSheB and Determination of its Complete Sequence

Plasmid pSheB, of the size of 81 kbp, was isolated from the *Shewanella* sp. O23S strain. In order to sequence the plasmid, plasmid pSheB was isolated from 200 ml of overnight culture of *Shewanella* sp. O23S by alkaline lysis method. Plasmid pSheB was sequenced by pyrosequencing method, using "shotgun" strategy on the GS FLX Titanium (454) sequencer (in the Oligo Pl. centre). For the construction of the DNA library, approx. 5 μg of pSheB DNA was used and reagent kits provided by the manufacturer were applied ((GS FLX Titanium Library Preparation Kit, Roche). The constructed library was sequenced and assembled using the software from the Newbler de novo assembler package (Roche). The obtained sequences were then assembled into contigs using Seqman software from Lasergene package (DNAStar). Annotation of the plasmid (identification of the open reading frames and determination of their potential functions) were performed using Artemis program and BLAST programs (from the NCBI database). The complete sequence of the plasmid has been shown in the SEQ ID NO: 1. Sequencing of the plasmid pSheB showed that it is a DNA particle of the size 81 591 bp and the GC-content of 44.04%. It comprises 87 open reading frames (ORF), which constitute 89.6% of the sequence of the plasmid. Table 1, below, features a detailed description of the identified ORFs within SEQ ID NO: 1.

TABLE 1

Determination of the potential coding sequences of the plasmid pSheB in reference to SEQ ID NO: 1.						
ORF #	Coding sequence (start-stop codon)*	Protein size (aa)	Predicted protein function	The greatest similarity (BLASTP program)		
				Identity (%)	Organism	GenBank number
1	1-813	270	Deoxyribonuclease I (EndA)	97 (233/240)	<i>Shewanella baltica</i> OS 195 (pS19501)	YP_001556993
2	826-1061	77	Hypothetical protein	100 (77/77)	<i>Shewanella baltica</i> BA175 (pSBAL17501)	YP_006018602
3	1454-3328	624	ParB like nuclease	99 (618/624)	<i>Shewanella baltica</i> BA175 (pSBAL17501)	AEG 13584
4	3453-3908	151	Hypothetical protein	99 (149/151)	<i>Shewanella baltica</i> OS195 (pS19501)	YP_001556996
5	4113-4397	94	Hypothetical protein	97 (91/94)	<i>Shewanella baltica</i> BA175 (pSBAL17501)	YP_006018598
6	4428-4649	73	Hypothetical protein	97 (71/73)	<i>Shewanella baltica</i> OS185 (pS18501)	YP_001355438
7	5014-5358	114	Hypothetical protein	100 (114/114)	<i>Shewanella baltica</i> OS195 (pS19502)	YP_001557025
8	5321-5944	207	Hypothetical protein	99 (204/207)	<i>Shewanella baltica</i> BA175 (pSBAL17501)	YP_006018595
9	5946-6347	133	Hypothetical protein	99 (132/133)	<i>Shewanella baltica</i> OS185 (pS18501)	YP_001355440
10	6545-6790	81	Toxin protein (HicA)	100 (81/81)	<i>Shewanella baltica</i> OS195 (pS19501)	YP_001557001
11	6790-7125	111	Antitoxin protein (HicB)	100 (111/111)	<i>Shewanella baltica</i> OS195 (pS19501)	YP_001557002
12	7423-7767	114	Hypothetical protein	98 (112/114)	<i>Shewanella baltica</i> OS195 (pS19502)	YP_001557025
13	7730-9058	442	Hypothetical protein	98 (4432/442)	<i>Shewanella baltica</i> OS625	EHC04198
14	9656-9967	103	Hypothetical protein	99 (102/103)	<i>Shewanella baltica</i> OS625	EHC04199
15	10260-10445c	61	Hypothetical protein	34 (13/38)	<i>Schefersomyces stipitis</i> CBS 6054	XP_001384086
16	10524-10883	119	Hypothetical protein	100 (119/119)	<i>Shewanella baltica</i> OS185 (pS18501)	YP_001355446
17	10904-16857c	1979	Conjugative transfer relaxase (TraI)	99 (1945/1969)	<i>Shewanella baltica</i> OS195 (pS19502)	YP_001557030
18	17109-19235c	708	Type IV conjugative transfer system coupling (TraD)	98 (694/708)	<i>Shewanella baltica</i> OS223 (pS22302)	YP_002360331
19	19908-22721c	937	Sex pilus assembly and mating pair (TraG)	95 (892/937)	<i>Shewanella baltica</i> OS223 (pS22302)	YP_002364250
20	22724-24112c	462	Type IV conjugative transfer system protein (TraH)	99 (460/462)	<i>Shewanella baltica</i> OS625	EHC04207
21	24296-24736c	146	Type-F conjugative transfer system pilin assembly thiol-disulfide isomerase (TrbB)	97 (141/146)	<i>Shewanella baltica</i> OS195 (pS19501)	YP_001556945
22	24750-25622c	290	Type-F conjugative transfer system pilin assembly protein (TraF)	99 (287/289)	<i>Shewanella baltica</i> OS 185 (pS18501)	YP_001355454
23	25622-27433c	603	Conjugal transfer mating pair stabilization protein (TraN)	87 (530/607)	<i>Shewanella baltica</i> OS223 (pS22302)	YP_002364256
24	27430-28170c	246	Type-F conjugative transfer system pilin assembly protein (TrbC)	100 (246/246)	<i>Shewanella baltica</i> OS195 (pS19501)	YP_001556948
25	28193-29200c	335	Sex pilus assembly and synthesis protein (TraU)	99 (332/335)	<i>Shewanella baltica</i> BA175 (pSBAL17501)	YP_006018570
26	29187-29891c	234	Type-F conjugative transfer system protein (TraW)	99 (232/234)	<i>Shewanella baltica</i> OS625	EHC04215
27	29888-30259c	123	Conjugal transfer protein (TrbI)	98 (120/123)	<i>Shewanella baltica</i> OS223 (pS22302)	YP_002360319
28	30261-32849c	862	Type-IV secretion system protein (TraC)	99 (856/862)	<i>Shewanella baltica</i> BA175 (pSBAL17502)	YP_006022865
29	32853-33296c	147	Type IV conjugative transfer system protein (TraV)	100 (147/147)	<i>Shewanella baltica</i> OS195 (pS19501)	YP_001556953
30	33329-34855c	508	Sex pilus assembly and synthesis protein (TraB)	98 (500/508)	<i>Shewanella baltica</i> OS678	YP_005280401
31	34852-35667c	271	Type-F conjugative transfer system secretin (TraK)	95 (258/271)	<i>Shewanella baltica</i> OS223 (pS22302)	YP_002364264

TABLE 1-continued

Determination of the potential coding sequences of the plasmid pSheB in reference to SEQ ID NO: 1.						
ORF #	Coding sequence (start-stop codon)*	Protein size (aa)	Predicted protein function	The greatest similarity (BLASTP program)		
				Identity (%)	Organism	GenBank number
32	35807-36232c	141	Type IV conjugative transfer system protein (TraE)	97 (118/122)	<i>Shewanella baltica</i> OS223 (pS22302)	YP_002364265
33	36273-36575c	100	Type IV conjugative transfer system protein (TraL)	94 (94/100)	<i>Shewanella baltica</i> OS195 (pS19501)	YP_001556957
34	36579-36953c	124	Type IV conjugative transfer system pilin (TraA)	87 (111/128)	<i>Shewanella baltica</i> BA175 (pSBAL17502)	YP_006022871
35	37017-37205c	62	Hypothetical protein	100 (62/62)	<i>Shewanella baltica</i> OS223 (pS22301)	YP_002360254
36	37307-37564c	85	Hypothetical protein with helix-turn-helix domain	100 (85/85)	<i>Shewanella baltica</i> OS195 (pS19501)	YP_001556959
37	37684-37866	60	Hypothetical protein	100 (60/60)	<i>Shewanella baltica</i> OS195 (pS19501)	YP_001556960
38	38139-38321	60	Hypothetical protein	32 (18/56)	<i>Staphylococcus aureus</i> subsp. <i>aureus</i> USA300_FPR3757	YP_494132
39	38308-39240	310	Hypothetical protein	97 (301/310)	<i>Shewanella baltica</i> OS195 (pS19503)	YP_001557103
40	39774-40076c	100	Pyridoxamine kinase family protein	27 (28/104)	<i>Megasphaera</i> <i>micronutriiformis</i> F0359	ZP_07757187
41	40161-40442c	93	Hypothetical protein	86 (75/87)	<i>Shewanella oneidensis</i> MR-1 (megaplasmid)	NP_720395
42	40448-40975	175	N-Acyltransferase superfamily protein	91 (160/175)	<i>Shewanella oneidensis</i> MR-1 (megaplasmid)	NP_720396
43	41417-41647c	76	Hypothetical protein	100 (74/74)	<i>Shewanella baltica</i> OS223 (pS22301)	YP_002360250
44	41822-43012	396	Plasmid partition protein (PatA)	99 (392/396)	<i>Shewanella baltica</i> BA175 (pSBAL17501)	AEG13539
45	43012-44169	385	Plasmid partition protein (ParB)	97 (347/356)	<i>Shewanella baltica</i> BA175 (pSBAL17501)	YP_006018554
46	44256-44402	48	Hypothetical protein			
47	44438-48298	1268	NTPase with transmembrane helices	25 (321/1284)	<i>Bacillus subtilis</i> subsp. <i>subtilis</i> str. 168	NP_389778
48	48454-49113	219	Hypothetical protein	60 (131/218)	<i>Methylophaga thiooxydans</i> DMSO10	ZP_05102952
49	49234-49623	129	Hypothetical protein	39 (49/125)	<i>Methylorovus glucosetrophus</i> SIP3-4	YP_003050141
50	49640-50545	301	Hypothetical protein	34 (102/300)	delta proteobacterium MLMS-1	ZP_01290544
51	50655-51623	322	Hypothetical protein	66 (212/322)	<i>Pseudoalteromonas arctica</i> A 37-1-2	ZP_10279906
52	51781-52386	201	Hypothetical protein	25 (44/176)	<i>Enterobacter cloacae</i> SCF1	YP_003941396
53	52593-53222c	209	Resolvase domain-containing protein (TnpR)	98 (204/209)	<i>Shewanella baltica</i> OS185 (pS18501)	YP_001355408
54	53345-53827	159	Transposase(TnpA)	72 (32/46)	<i>Vibrio furnissii</i> CIP 102972	EEX38686
55	53836-54123	95	Hypothetical protein	83 (79/95)	<i>Marinomonas</i> sp. MWYL1	ABR70068
56	54532-54840	102	ArsR family transcriptional regulator	94 (96/102)	<i>Shewanella</i> sp. ANA-3	YP_869986
57	54901-55143	80	Thioredoxin - redox-active disulfide protein 2	93 (74/80)	<i>Shewanella</i> sp. ANA-3	YP_869985
58	55162-55695	177	Hypothetical protein	95 (169/177)	<i>Shewanella</i> sp. ANA-3	YP_869984
59	55706-56386	226	Cytochrome c biogenesis protein	99 (225/226)	<i>Shewanella</i> sp. ANA-3	YP_869983
60	56554-57555	333	RND family efflux transporter MFP subunit	93 (311/333)	<i>Shewanella</i> sp. ANA-3	YP_869982
61	57552-60625	1023	Acriflavin resistance protein (AcrB)	98 (1002/1023)	<i>Shewanella</i> sp. ANA-3	YP_869981
62	60743-61177c	144	Arsenate reductase (ArsC)	94 (131/140)	<i>Shewanella</i> sp. ANA-3	YP_869980
63	61265-62512c	415	Arsenic pump membrane protein (ArsB)	98 (407/414)	<i>Shewanella</i> sp. W3-18-1	YP_964320
64	62613-64385c	590	Arsenite-activated ATPase (ArsA)	92 (543/590)	<i>Shewanella</i> sp. W3-18-1	YP_964319
65	64420-64782c	120	Arsenical resistance operon transacting repressor (ArsD)	85 (102/120)	<i>Shewanella</i> sp. W3-18-1	YP_964318

TABLE 1-continued

Determination of the potential coding sequences of the plasmid pSheB in reference to SEQ ID NO: 1.						
ORF #	Coding sequence (start-stop codon)*	Protein size (aa)	Predicted protein function	The greatest similarity (BLASTP program)		
				Identity (%)	Organism	GenBank number
66	65155-67719	854	Respiratory arsenate reductase, Mo binding subunit (ArrA)	96 (818/854)	<i>Shewanella</i> sp. ANA-3	YP_869976
67	67731-68435	234	Respiratory arsenate reductase, FeS subunit (ArrB)	97 (228/234)	<i>Shewanella</i> sp. ANA-3	YP_869975
68	68506-69936c	476	Glutathione synthase	86 (409/476)	<i>Shewanella putrefaciens</i> CN-32	YP_001182743
69	70092-71144c	350	Permease	76 (271/358)	<i>Shewanella putrefaciens</i> CN-32	YP_001185328
70	71369-71656	95	ArsR family transcriptional regulator	92 (87/95)	<i>Shewanella</i> sp. ANA-3	YP_869972
71	71755-72171	138	Transcriptional regulator - tyrosine phosphatase (ArsR)	93 (127/138)	<i>Shewanella</i> sp. ANA-3	YP_869971
72	72164-72511	115	ArsR family transcriptional regulator	97 (112/15)	<i>Shewanella</i> sp. ANA-3	YP_869970
73	72590-73102	170	Protein tyrosine phosphatase (ArsC2)	84 (141/167)	<i>Shewanella</i> sp. ANA-3	YP_869969
74	73195-74192	332	Permease	96 (318/332)	<i>Shewanella</i> sp. ANA-3	YP_869968
75	74204-74440	78	Redox-active disulfide protein 2	97 (76/78)	<i>Shewanella putrefaciens</i> CN-32	YP_001185322
76	74448-74939	163	Dual specificity protein phosphatase	86 (140/163)	<i>Shewanella putrefaciens</i> CN-32	YP_001185321
77	74996-76006	336	Glyceraldehyde-3-phosphate dehydrogenase	99 (332/336)	<i>Shewanella</i> sp. W3-18-1	YP_964294
78	76012-77250	412	Major facilitator superfamily protein	99 (409/412)	<i>Shewanella putrefaciens</i> 200	YP_006009170
79	77357-77656	99	Hypothetical protein	92 (91/99)	<i>Shewanella putrefaciens</i> CN-32	YP_001185318
80	77964-78146	60	Hypothetical protein	32 (13/41)	<i>Streptococcus pneumoniae</i>	CAI34125
81	78249-78683	144	Peptidase S24/S26A/S26B	97 (139/144)	<i>Shewanella baltica</i> BA175 (pSBAL17501)	YP_006018610
82	78671-79924	417	UMUC domain-containing protein DNA-repair protein	93 (389/417)	<i>Shewanella baltica</i> BA175 (pSBAL17501)	YP_006018609
83	79975-80238	87	Vitamin B12 dependent methionine synthase	30 (24/81)	<i>Coprococcus comes</i> ATCC 27758	ZP_03798710
84	80344-80994	216	Hypothetical protein	68 (150/219)	<i>Aeromonas hydrophila</i>	YP_002995563
85	81035-81253	72	Hypothetical protein	93 (67/72)	<i>Shewanella baltica</i> OS117 (pSBAL11701)	YP_006035206
86	81273-81591	106	Hypothetical protein	96 (102/106)	<i>Shewanella baltica</i> BA175 (pSBAL17501)	YP_006018603

*The numbers in the coding sequence correspond to the nucleotide numbers in SEQ ID NO: 1.

Example 2

Construction of the Plasmid-Less Strain and Functional Analysis of the Plasmid pSheB

In order to show, that the plasmid pSheB and the gene module coding for potential proteins located within it, are involved in the resistance to arsenic and dissimilatory arsenates reduction, a plasmid-less derivative of the *Shewanella* sp. O23S strain was constructed and its functional analysis was carried out. As the stress factor stimulating the mechanisms of plasmid removal from the cells of the host, ethidium bromide (EtBr) solution at a final concentration of 5 µM was used.

Overnight culture of the wild-type *Shewanella* sp. O23S strain, carried out in LB medium supplemented with 5 mM sodium arsenate was passaged to LB medium with 5 µM EtBr. The optical density (OD) of the culture at the beginning of the experiment was OD=0.1, and the culture was carried out for 24 hours at 22° C. with shaking (160 rpm). After 24 hours of incubation, culture dilutions 10⁻⁴, 10⁻⁶, 10⁻⁸, 10⁻¹⁰ were prepared, respectively, and 100 µl of each of them were plated on LB

medium solidified with agar. Subsequently, 96 colonies were randomly selected and passaged by replica plating to:

- (i) solid LB medium,
- (ii) solid LB medium supplemented with sodium arsenate (50 mM), and
- (iii) liquid minimal R1-R2 medium (R1 salt: NaCl—1.17 g/l; KCl—0.3 g/l; NH₄Cl—0.15 g/l; MgCl₂·6H₂O—0.41 g/l; CaCl₂·2H₂O—0.05 g/l and R2 salt (KH₂PO₄—0.17 g/l; NaHCO₃—2.0 g/l; Na₂SO₄·10 H₂O—0.07 g/l mixed in a ratio 1:1), enriched with sodium lactate at a final concentration of 5 mM, salts according to Tuovinen (2 ml/l) (pH 6.0; with the composition: Na₂EDTA 50 g/l; ZnSO₄·7H₂O 11 g/l; MnCl₂·7H₂O 5.5 g/l; FeSO₄·7H₂O 2.5 g/l; (NH₄)₆Mo₇O₂₄·4H₂O 5 g/l; CuSO₄·5H₂O 2 g/l; CoCl₂·6H₂O 0.5 g/l; NaOH 11 g/l) (Tuovinen et al., 1973), yeast extract at a final concentration of 0.004% and 2.5 mM sodium arsenate.

The cultures in solid media were incubated for 72 hours, whereas the cultures in the liquid medium were carried out in

15

200 µl in 96-well titration plates in Anaerocult® (Merck) containers, which provide anaerobic conditions for 168 hours.

The cultures carried out in the liquid medium were aimed to determine the abilities of the selected strains to dissimilatory reduce arsenates. After 5 days of incubation under anaerobic conditions, 100 µl of 0.1 M solution of silver nitrate were added to the cultures. The result of the reaction between AgNO_3 and As (III) or As (V) is the formation of a coloured precipitate. A brown precipitate indicates the presence of Ag_3AsO_4 (silver orthoarsenate), while a yellow precipitate indicates the presence of Ag_3AsO_3 (silver arsenite). In case of testing for the ability to reduce arsenates, the presence of a yellow precipitate indicates that As(V) was reduced to As(III).

The cultures carried out in solid LB medium enriched with sodium arsenate were aimed to determine the resistance to As(V). In turn, the cultures carried out in solid LB medium (not enriched with additional substances) were aimed to secure the potential mutants (positive control). All the strains, which have grown in LB medium, and were not capable of growth in LB medium enriched with As(V) and in minimal R1-R2 medium containing arsenate (the final electron acceptor) and lactate (electron donor), have been designated as potential mutants, deprived of the plasmid pSheB.

In order to verify the selected consortia, the plasmid profile of the wild-type strain (*Shewanella* sp. O23S) and the potential plasmid-less mutants was checked. The plasmid DNA was isolated by alkaline lysis method, and electrophoretic analysis (0.8% agarose gel) was conducted. The comparison of the plasmid patterns of the selected strains allowed for the identification of the plasmid-less mutants. An additional confirmation of the absence of the plasmid pSheB in the cells of the constructed mutants was PCR analysis. PCR reaction was carried out in the genomes of the potential mutants, using the following primers:

endA-L GCTGTTGCTTCCAATACGAC (SEQ ID NO: 2)
and
endA-R GGCCTGCGACTTACTCATC (SEQ ID NO: 3)

Primer endA-L respectively corresponds to the position of the nucleotide 127, while endA-R—position 679 of the plasmid pSheB in reference to SEQ ID NO: 1. The strains (potential mutants) which gave a negative result in the PCR reaction, using the primers described above, were selected for further analysis.

The next step of the verification of the strains not-possessing the plasmid pSheB was their functional analysis. It was checked again whether the selected strains are capable of dissimilatory arsenate reduction. For this purpose, tests were carried out in minimal R1-R2 medium enriched with sodium arsenate and sodium lactate, and applying the test with 0.1 M solution of silver nitrate. The strains, which were not capable of reduction, were susceptible to As(V) and As(III), and did not possess the plasmid pSheB, turned out to be a proof that the plasmid pSheB determines the capability of arsenate respiration and resistance to arsenic. FIG. 2. shows a graph illustrating the kinetics of arsenate reduction of the wild-type strain and the mutant deprived of the plasmid pSheB. The strain deprived of the plasmid was not capable of growth in minimal medium supplemented with 2.5 mM sodium arsenate, thus it was unable to reduce As(V) to As(III). In this way it was shown that the genetic information contained in the plasmid pSheB (SEQ ID NO: 1) determines the acquisition of the capability of growth under anaerobic conditions, using

16

arsenates as the final electron acceptor, thus the capability of dissimilatory arsenate reduction.

Example 3

Construction of the Vector Carrying a Gene Module Coding for the Proteins Involved in Dissimilatory Arsenate Reduction

In order to demonstrate, which genes located on the plasmid pSheB encode proteins responsible for dissimilatory arsenate reduction, the arr module, which may comprise i.a. genes for dissimilatory arsenate reductase arrAB, was cloned into the vector pBBR1-MCS2 (Km'), in the *Escherichia coli* TOP10 strain, and its functionality was tested.

In order to clone the arr module, amplification of a DNA fragment of the size 8634 bp (which may comprise the region from position 63978 to 72599 in the genome of pSheB) was performed on a DNA template of the plasmid pSheB, isolated by alkaline lysis. For PCR reaction, the following oligonucleotides were used as primers:

She_Mph1103F: GAAATCTTGCAGTAGCGATGCATC (SEQ ID NO: 4) [position in the genome of the plasmid pSheB: 63978-64001; the underlined sequence is the restriction site recognized by the enzyme Mph11031 (NsI)], and

She_XmaJR: GTTGTTCCTAGGCTGGTGCCATATCACACCTCTAG (SEQ ID NO: 5) [position in the genome of the plasmid pSheB: 72578-72599; the sequence written in italic is an added sequence; the underlined site is recognized by the restriction enzyme XmaJI]. For the amplification, Phusion® High-Fidelity DNA Polymerase (Thermo Scientific) was used.

The obtained PCR product (8634bp) was cloned into a plasmid vector: pBBR1MCS-2 (Km') [Kovach et al., 1995] digested (linearized) with SmaI. The ligation mixture of the PCR product and the vector pBBR1MCS2 digested with the enzyme SmaI was introduced, by means of chemical transformation, using the calcium-rubidium method according to Kushner (1978), into the cells of *Escherichia coli* Top10 strain [mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 recA1 araD139 Δ(ara-leu)7697 galU galK rpsL endA1 nupG]. Complete LB medium with kanamycin (30 µg/ml), IPTG (0.5 µg), and X-gal (40 µg/ml) was used as the selection medium.

From the pool of the obtained transformants (white colonies resistant to kanamycin) strains that were harbouring a plasmid of the appropriate size: 13778 bp (pBBR1MCS2-5144bp+arr module—8634 bp). The presence of the constructed plasmid was confirmed by restriction analysis (digestion with the enzymes XmaJI and Mph11031), electrophoretic analysis and sequencing. The *Escherichia coli* MR1 strain (derivative of the *E. coli* TOP10 strain) harbouring the plasmid pARR1A (derivative of pBBR1MCS2 with cloned arr module), was selected for further analysis.

In order to verify the functionality of the constructed plasmid pARR1A, phenotypic analysis of the *E. coli* MR1 strain was carried out in minimal R1-R2 medium enriched with 2 mM sodium arsenate and 5 mM sodium lactate and supplemented with 0.004% yeast extract. The culture was carried out for 120 h under anaerobic conditions (in $\text{CO}_2:\text{N}_2$ atmosphere) at 37°C. After five days of culturing, the test with 0.1 M solution of silver nitrate was carried out, which showed that the investigated strain has reduced arsenates to arsenites. A confirmation of the reduction of arsenates to arsenites by the *E. coli* MR1 (pARR1A) strain, was the qualitative analysis of arsenic speciation by HPLC. The conducted analysis showed that after 120 h of incubation of the *E. coli* MR1

17

strain, arsenites were identified in the medium. In the control reaction with the *E. coli* TOP10 strain (without the plasmid), reduction of As(V) to As(III) was not observed.

The obtained results showed that the introduction of the genes of the arr module, originating from the genome of the plasmid pSheB, on the vector pBBR1MCS into the *E. coli* TOP10 strain, leads to the acquisition of the capability of dissimilatory arsenate reduction. The module introduced to *E. coli* TOP10 corresponded to the fragment of the nucleotide sequence from 63978 to 72599 of the plasmid pSheB, having the sequence shown in SEQ ID NR: 1. In this way it was demonstrated that this is the fragment of the plasmid responsible for the capability of dissimilatory arsenate reduction. It was also demonstrated, that a bacterium other than that, from which the plasmid pSheB originates, that is another species of bacteria or a bacterial strain, into which the fragment of the sequence which may comprise the fragment of the nucleotide sequence from 63978 to 72599 of the plasmid pSheB, having the sequence shown in SEQ ID NO: 1 has been introduced, acquires the ability to dissimilatory reduce arsenates.

Example 4

Removal of Arsenic from Copper Minerals Using the *Shewanella* sp. O23S Strain

In order to demonstrate that a strain which may comprise the plasmid pSheB (SEQ ID NO: 1), for example the *Shewanella* sp. O23S strain, harbouring the plasmid pSheB, can be used in biometallurgy, in the processes of selective removal of arsenic from mineral resources, an experiment was carried out, using two types of minerals: copper-bearing bituminous shale of the "Kupferschiefer" type and flotation tailings from the first series, obtained from the Mining Plant "Lubin" (KGHM, Poland). The average arsenic content in the copper-bearing bituminous shale designated as "Shales" was 1000-3000 mg/kg of the shale, whereas the arsenic content in the flotation tailings designated as "Middlings" was 250-350 mg/kg of the tailings. Apart from arsenic, both types of minerals contain Cu in the range of 35000-110000 mg/kg of dry mass, and other precious metals, e.g. Co—500-2500 mg/kg, Zn—15-2800 mg/kg, Ni—250-500 mg/kg.

The experiment of arsenic removal from the mineral resources described above was conducted in R1-R2 medium enriched with 5 mM sodium lactate as the source of carbon and energy, Tuovinen salts (2 ml/l) and yeast extract at a final concentration of 0.004%. The mineral substrate (shales or flotation tailings) having the fraction size of 125-250 mm, was added to a final concentration of 10%. The cultures were carried out in 100-ml bottles under anaerobic conditions, in the atmosphere of the mixture of gases N₂:CO₂ (4:1) for 21 days at 22° C. with shaking (160 rpm). In order to obtain an inoculum of the *Shewanella* sp. O23S strain, overnight cultures were started in liquid complete LB medium. The overnight culture of the O23S strain was centrifuged several times and flushed with saline solution, and then passaged to an appropriate medium, to obtain a density of approximately 10⁶ cells/ml. Sterile medium supplemented with an appropriate substrate, not inoculated with bacteria, was used as the control.

At the beginning of the experiment, and every 7 days, samples for the determination of copper and arsenic content were collected, which was performed using atomic absorption spectroscopy—flame technique; (AA Solaar M6 Spectrometer, TJA Solutions, UK).

The conducted analyses showed, that the highest concentration of arsenic in culture fluids in medium enriched with

18

flotation tailings was noted after 14 days of incubation (324.75 µg/l) (FIG. 3A). In a further week of incubation, the concentration of arsenic has dropped (192.75 mg/l), which may be associated with precipitation of arsenic in the form of the secondary minerals, in the control samples, concentration of arsenic was at much lower level (it did not exceed 153.75 µg/l), and reflected the chemical processes of leaching. In turn, in culture fluids, in medium enriched with bituminous shales, the highest concentration of arsenic was noted after 21 days of the experiment (678.25 µg/l). At the same time, in the control sample, an approx. 14-times lower concentration of arsenic (47.25 µg/l) was noted. These results indicate that the *Shewanella* sp. O23S strain with the plasmid pSheB (SEQ ID NO: 1) is capable of releasing arsenic from copper minerals. It was therefore extremely important to verify how the release of arsenic would influence copper mobilization. In case of copper content analysis, no significant differences were observed between the Cu content in culture fluids and the control sample. At the beginning of the experiment it was noted that part of the copper was washed out of the minerals as a result of the chemical (stimulated by the medium components) dissolution of copper minerals (FIG. 4). The release of copper occurred with an extremely low efficiency, and Cu concentrations in culture fluids and the control samples did not exceed 3.5 ppm. Furthermore, with the passing of time, the concentration of copper was decreasing (after 21 days of cultivation, the concentration of Cu was below 0.5 ppm (FIG. 4.), which was probably associated with the chemical precipitation of copper.

The conducted experiment allowed to demonstrate that the strain which may comprise the plasmid pSheB (SEQ ID NO: 1), i.e. the *Shewanella* sp. O23S strain, deposited as KKP2045p, removes arsenic from copper-bearing bituminous shales, as well as from flotation tailings in a specific way, without simultaneous mobilization of copper from copper deposits.

Example 5

Mobilization of Arsenic from the Soils Contaminated with Arsenic Using the *Shewanella* sp. O23S Strain

In order to demonstrate that a strain harbouring the plasmid pSheB, having the SEQ ID NO: 1 or its functional derivative, such as *Shewanella* sp. O23S, harbouring the plasmid pSheB, can be used in bioremediation, in the processes of selective arsenic removal from the soils contaminated with arsenic, an experiment using the soil originating from the vicinity of a gold mine in Zloty Stok was conducted. The average arsenic content in the soil was 17955.7 mg/kg of soil.

The experiment of arsenic removal from the soil was conducted in R1-R2 medium enriched with 5 mM sodium lactate (carbon and energy source), Tuovinen* salts (2 ml) and yeast extract at a final concentration of 0.004%. The fragmented sample (having the fraction size of <3 mm) was added to a final concentration of 10%. The cultures were carried out in 100-ml bottles under anaerobic conditions in the atmosphere of the mixture of gases N₂:CO₂ (4:1) for 21 days at 22° C. with shaking (160 rpm). In order to obtain an inoculum of the *Shewanella* sp. O23S strain, overnight cultures were started in liquid complete LB medium. The overnight culture of the O23S strain was centrifuged several times and flushed with saline solution, and then passaged to an appropriate medium to obtain a density of approximately 10⁶ cells/ml. Sterile medium with the addition of the soil, not inoculated with bacteria, was used as the control.

At the beginning of the experiment, and every 7 days, samples were collected for: (i) determination of the arsenic content (determination by atomic absorption spectroscopy—flame technique) in culture fluids (ii) monitoring the growth of bacteria, by the analysis of the number of colony forming units (cfu) (plating on solid LB medium enriched with 5 mM sodium arsenate). The conducted analyses showed, that the highest concentration of arsenic in culture fluids was noted after 14 days of incubation (166.53 mg/l), whereas after a further 7 days of incubation, concentration of arsenic has decreased slightly (123.03 mg/l), which may be associated with the precipitation of arsenic in the form of secondary minerals. In the control samples, concentration of arsenic was at much lower level (it did not exceed 25 mg/l), and reflected the chemical processes of leaching. It was demonstrated that a strain which may comprise the plasmid pSheB, having the SEQ IN NO:1, i.e. the *Shewanella* sp. O23S strain, removes arsenic from the soils contaminated with this element and can be used in bioremediation, e.g. in the recultivation of arsenic contaminated soils or the removal of arsenic from other contaminated environments.

Example 6

Analysis of Arsenic Accumulation by the *Shewanella* sp. O23S Strain

The growth experiment and the dissimilatory arsenate reduction performance analysis, carried out in R1-R2 medium (FIG. 2) revealed that a strain harbouring the plasmid pSheB, having the SEQ ID NO:1 or its functional derivative, such as *Shewanella* sp. O23S, completely reduces arsenates to arsenites, which are completely removed out of the cell. In order to confirm that the *Shewanella* sp. O23S strain is not capable of accumulating arsenic inside the cells, an additional 24-hour growth experiment was carried out in liquid LB medium. The LB medium was enriched with a solution of sodium arsenites (Na_3AsO_2) or sodium arsenates (Na_2HAsO_4) having a final concentration of 2 mM. The cultures were carried out in a volume of 50 ml under aerobic conditions at 22° C. with shaking (160 rpm). Sterile LB medium with solutions of the investigated salts was used as the control. After 24 hours of incubation, cultures were centrifuged and the arsenic content in culture fluids and biomass (bacterial pellet) was determined by atomic absorption spectroscopy (AA Solaar M6 Spectrometer, TJA Solutions, UK).

The conducted analyses did not confirm the growth experiments in R1-R2 medium and revealed that arsenic compounds may be partially accumulated inside the cells of the *Shewanella* sp. O23S strain. The efficiency of this process, however is very low, less than 3% for As(III) compounds, and less than 8% for As(V) compounds. The obtain results are shown in Table 2.

TABLE 2

Comparison of the ability to accumulate arsenic by the *Shewanella* sp. O23S strain, cultured in minimal medium (R1-R2) and complete medium (LB).

	LB medium		R1-R2 medium	
	As(III)	As(V)	As(III)	As(V)
Arsenic content in culture fluids [ppm]	144.8119	163.8377	204.8073	N/D
Arsenic content in the biomass [ppm]	5.809583	13.97905	4.1078	N/D

TABLE 2-continued

Comparison of the ability to accumulate arsenic by the *Shewanella* sp. O23S strain, cultured in minimal medium (R1-R2) and complete medium (LB).

	LB medium		R1-R2 medium	
	As(III)	As(V)	As(III)	As(V)
Accumulation efficiency [%]	3.857076	7.861492	1.9636	N/D
N/D-not determined				

Example 7

Analysis of the Production of Volatile Arsenic Compounds by the *Shewanella* sp. O23S Strain

In order to verify whether the *Shewanella* sp. O23S strain which may comprise the plasmid pSheB is capable of producing volatile arsenic compounds, an anaerobic (in $\text{N}_2:\text{CO}_2$ atmosphere; 4:1) culture in minimal R1-R2 medium, enriched with 5 mM sodium lactate and 2.5 mM sodium arsenite, was carried out. After 5 days of incubation at 22° C. gas samples were collected for the chemical composition analysis in a gas chromatograph GC-MS and GC-AED. The conducted analyses revealed that the only volatile compound produced by the *Shewanella* sp. O23S strain with the plasmid pSheB is dimethyl monosulfide (DMS). The production of volatile arsenic compounds was not observed.

Example 8

Precipitation of As_2S_3 (Arsenic(III) Sulfide) by the *Shewanella* sp. O23S Strain in Medium Containing As (V) and Thiosulfate

In order to demonstrate that a strain harbouring the plasmid pSheB, having the SEQ ID NO: 1 or its functional derivative, such as *Shewanella* sp. O23S, harbouring the plasmid pSheB can be used in bioremediation, in the processes of selective arsenic removal from the soils contaminated with arsenic, an experiment, in which arsenic was precipitated from the medium in the form of arsenic (III) sulfide (As_2S_3) was carried out. For this purpose, an experiment of arsenic removal from the soil was conducted in R1-R2 medium enriched with 5 mM sodium lactate (as the source of carbon and energy), Tuovinen salts (2 ml/l), and 2.5 mM sodium arsenite and 5 mM sodium thiosulfate as the final electron acceptors. The medium was inoculated to obtain the initial culture density of 10^6 cells/ml. The cultures were carried out under anaerobic conditions, at room temperature. The arsenic content in the solution was measured at time T0, and on days 1, 2, 3, 4, and 5, 7, 14 and 21 of the culture. The appearance of a yellow-orange arsenic (III) sulfide (As_2S_3) precipitate was also observed. After 7 days, a loss of 47.8% of arsenic in the solution, and the appearance of a characteristic precipitate, insoluble in water, were recorded. After 21 days of cultivation, the loss of arsenic reached 82.3% (FIG. 6). In the controls containing a sterile medium, sodium thiosulfate or sulfide, or sulfites (IV) or sulfates (VI), as well as arsenites (III) or arsenates (V), neither the loss of arsenic in the solution, nor the appearance of arsenic sulfide precipitates were observed. It was hereby demonstrated that the strain which may comprise the plasmid pSheB (SEQ ID NO: 1), i.e. the *Shewanella* sp. O23S strain deposited as KKP2045p, is capable, in the presence of min-

eral sulfur compounds, of arsenic removal from solutions (e.g. of soil) in the form of arsenic (III) sulfide, and thus, of its immobilization.

Example 9

Capability of Dissimilatory Arsenate Reduction by the *Shewanella* sp. O23S Strain Under Slightly Acidic Conditions

In order to verify whether the *Shewanella* sp. O23S strain is capable of dissimilatory arsenate reduction under slightly acidic conditions, culture was carried out in minimal R1-R2 medium having a pH 4, enriched with 2.5 mM sodium arsenate, 5 mM sodium lactate and 0.004% yeast extract. The culture was carried out for 168 h at room temperature. An analogous experiment in medium having a pH 8 was carried out as the control. The conducted research revealed, that in the most optimal conditions, in medium having a pH 8, *Shewanella* sp. O23S has completely reduced As(V) present in the medium within 4 h (FIG. 7A). On the other hand, in medium having a pH 4, *Shewanella* sp. O23S metabolism, and the kinetics of the process of reduction of arsenates to arsenites have been slowed down, and the complete reduction of As(V) occurred after 48 h (FIG. 7B). Nevertheless, the obtained results revealed that the *Shewanella* sp. O23S strain is capable of dissimilatory arsenate reduction under slightly acidic conditions.

Example 10

Resistance to Heavy Metals of the *Shewanella* sp. O23S Strain

In order to use a strain in mobilization of arsenic from polymetallic deposits, copper concentrates or flotation tailings, the used strain must be characterized by a high resistance to heavy metals. In accordance with the above, *Shewanella* sp. O23S capability of tolerating the presence of heavy metals, and capability of reducing As(V) to As(III) in the presence of heavy metals was verified.

In order to verify the range of tolerance to the presence of heavy metals, the *Shewanella* sp. O23S strain was cultured in liquid complete LB medium supplemented with the appropriate heavy metals solutions (Table 3):

TABLE 3

Metals and their compounds used for the determination of the minimal inhibitory concentration (MIC).			
Analysed metal	Chemical compound	Concentration [mM]	Concentration [mg/l]
As(III)	NaAsO ₂	0.5-5	37.5-375
As(V)	Na ₂ HA ₂ O ₄	1-600	75-45000
Cr (III)	Cr ₂ (SO ₄) ₃ •18H ₂ O	2-12	104-624
Zn (II)	ZnSO ₄ •7H ₂ O	1-6	65.5-393
Se (VI)	Na ₂ SeO ₄ •10H ₂ O	1-20	79-1580
Cu (II)	CuSO ₄ •5H ₂ O	1-6	63.5-380
Co (II)	CoSO ₄ •7H ₂ O	1-6	59-354
Mn (II)	MnSO ₄ •H ₂ O	1-20	55-330
V (V)	NaVO ₃	1-20	51-1020
Cd (II)	CdSO ₄ •8H ₂ O	1-6	112-672

The overnight culture of *Shewanella* sp. O23S was passed to LB medium enriched with the appropriate metal or metalloid compound (to obtain a density of approximately 10⁶ cfu/ml) and was incubated for 24 hours at 22° C. Subsequently, OD₆₀₀ measurements were carried out and the values

of minimal inhibitory concentration (MIC), which is defined as the lowest Meⁿ⁺ concentration that completely inhibits the growth of bacteria, were determined. In order to verify whether the presence of heavy metals inhibits the capability of reducing arsenates to arsenites, a test was carried out in minimal R1-R2 medium enriched with 2.5 mM sodium arsenate, 5 mM sodium lactate, 0.004% yeast extract and the addition of the appropriate heavy metals solutions.

The conducted experiments revealed that the *Shewanella* sp. O23S strain, apart from the resistance to As(III) and As(V) is also resistant to Cu, Cd, Cr, Co, Mn, Zn, Se, V (Table 4). Furthermore, it is also capable of dissimilatory arsenate reduction in the presence of heavy metals. Only in the presence of iron (III) the capability of dissimilatory reduction was not demonstrated, because this strain is capable of Fe(III) reduction and using it as the final electron acceptor (Table 4).

TABLE 4

Element	MIC value [mM]	Reduction of arsenic in the presence of the metal [mM]
Cd(II)	1	1
Co(II)	2	2
Cr(III)	5	3
Cu(II)	3	2
Fe(III)	5	—
Mn(II)	>20	>5
Se(VI)	>20	5
Zn(II)	3	2
V(V)	>20	>5

LITERATURE CITED IN THE DESCRIPTION, IS IN ITS ENTIRETY INCORPORATED HEREIN AS REFERENCES

- Bultreys A. and I. Gheysen. 2000. Production and comparison of peptide siderophores from strains of distantly related pathovars of *Pseudomonas syringae* and *Pseudomonas viridisflava* LMG 2352. Appl. Environ. Microbiol. 66:325-31.
- Drewniak L. 2009. Characterization of arsenic bacteria isolated from Zloty Stok gold mine. PhD thesis. The Faculty of Biology, University of Warsaw.
- Drewniak L., Matlakowska R., Rewerski B. and Sklodowska A. 2010. Arsenic release from gold mine rocks mediated by the activity of indigenous bacteria. Hydrometallurgy 104 (3-4): 437-442
- Fornasiero D., Fullston C. L. and Ralston J. 2001. Separation of enargite and tennantite from non-arsenic copper sulfide minerals by selective oxidation or dissolution. International Journal of Mineral Processing, 61 (2): 109-119
- Guo H. and Yen W. T. 2005. Selective flotation of enargite from chalcopyrite by electrochemical control, Miner Eng. 18(6):605-612.
- Kovach M. E., Elzer P. H., Hill D. S., Robertson G. T., Farris M. A., Roop R. M., Peterson K. M. 1995. Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. Gene, 166:175-6.
- Kushner S. R. 1978. An improved method for transformation of *E. coli* with ColE1 derived plasmids, str. 17-23. Boyer H. B. i S. Nicosia (ed.). Genetic engineering. Elsevier/North-Holland, Amsterdam

- Mantur A., Rajpert L., Rewerski B., Ruszkowski D., Skłodowska A. and Drewniak L. 2011. New dissimilatory arsenate reducers—isolation, characteristic and potential application in biometallurgy. Prezentacja na konferencji BioMicroWorld 2011, Malaga Hiszpania
- Newman D. K., Ahmann D. and Morel F. M. 1998. A brief review of microbial arsenate respiration. *Geomicrobiol. J.* 15:255-268
- Olson G. J., Brierley J. A. and Brierley C. L. 2003. Bioleaching review part B: progress in bioleaching: applications of microbial processes by the minerals industries. *Appl Microbiol Biotechnol.* 63(3):249-257
- Piret N. L. 1999. The removal and safe disposal of arsenic in copper processing. *JOM*, 51 (9) 16-17
- Rawlings D. E. and Johnson D. B. 2007. A Review: The microbiology of biomining: development and optimization of mineral-oxidizing microbial consortia. *Microbiol.* 153: 315-324
- Robins R. G. 1985. The aqueous chemistry of arsenic in relation to hydrometallurgical processes. Proceedings of the 15th Annual CIM Hydrometallurgical Meeting, Vancouver, Canada, pp. 11-126.
- Sambrook J. and Russell D. W. 2001 Molecular cloning: A laboratory manual. Cold Spring Harbor Laboratory Press, New York.
- Schwyn B. and J. B. Nelands. 1987. Universal chemical assay for the detection and determination of siderophores. *Anal. Biochem.* 160:47-56.
- Senior G. D. Guy P. J. and Bruckard W. J. 2006. The Selective Flotation of Enargite from Other Copper Minerals—A Single Mineral Study in Relation to Beneficiation of the Tampakan Deposit in the Philippines. *Internation Journal of Mineral Processing*, 81: 15-26
- Skłodowska A. and Matlakowska R. 2007. Bioleaching of metals in neutral and slightly alkaline environment In Microbial Processing of Metal Sulfides Edts.: Edgardo R. Donati Wolfgang Sand Published by Springer, Dordrecht, The Netherlands, ISBN-10 1-4020-5588-9 (HB); ISBN-13 978-1-4020-5588-1 (HB) pp. 121-130
- Tuovinen O. H. and D. P. Kelly. 1973. Studies on the growth of *Thiobacillus ferrooxidans*. I. Use of membrane filters and ferrous iron agar to determine viable numbers, and comparison with 14 CO₂—fixation and iron oxidation as measures of growth. *Arch. Mikrobiol.* 88:285-98.
- Xia L., Yin C., Dai S., Qin G., Chen X. and Liu J. 2010. Bioleaching of chalcopyrite concentrate using *Leptospirillum ferrophilum*, *Acidithiobacillus ferrooxidans* and *Acidithiobacillus thiooxidans* in a continuous bubble column reactor. *J Ind Microbiol Biotechnol.* 37 (3): 289-295
- Xia L., Dai S., Yin C., Hu Y., Liu J., and Qiu G. 2009. Comparison of bioleaching behaviors of different compositional sphalerite using *Leptospirillum ferrophilum*, *Acidithiobacillus ferrooxidans* and *Acidithiobacillus caldus*. *J Ind Microbiol Biotechnol.* 36(6):845-85
- Zobrist J., Dowdle P. R., Davis J. A. and Oremland R. S. 2000. Mobilization of Arsenite by Dissimilatory Reduction of Adsorbed Arsenate. *Environ. Sci. Technol.* 34:4747-4753.
- The invention is further described by the following numbered paragraphs:
1. An isolated plasmid comprising the fragment of the nucleotide sequence from 63978 to 72599 of the plasmid pSheB, having the sequence shown in SEQ ID NO: 1 or its functional derivative.
 2. An isolated plasmid pSheB, having the sequence shown in SEQ ID NO: 1 or its functional derivative.
 3. A bacterial strain comprising the plasmid defined in paragraph 1 or 2, or comprising the nucleotide sequence

- comprising the nucleotides from 63978 to 72599 of SEQ ID NO: 1 or its functional derivative.
4. The bacterial strain according to paragraph 3, characterised in that it is the *Shewanella* sp. O23S strain deposited in the IAFB Collection of Industrial Microorganisms in Warsaw, under the deposit number KKP 2045p.
5. A composition comprising the isolated plasmid defined in paragraphs 1-2 or the bacterial strain defined in paragraphs 3-4 or combination thereof.
6. Use of the bacterial strain defined in paragraphs 3-4 and/or the composition defined in paragraph 5 for the selective removal of arsenic from mineral resources, raw materials industry waste or soil.
7. The use according to paragraph 6, characterised in that the mineral resources are polymetallic copper deposits.
8. A method for selective removal of arsenic from mineral resources, raw materials industry waste or soil, wherein the step of dissimilatory arsenate reduction is carried out with the use of the bacterial strain defined in paragraphs 3-4 and/or the composition defined in paragraph 5.
9. The method, according to paragraph 8, wherein the step of dissimilatory arsenate reduction is carried out under neutral or slightly alkaline conditions.
10. The method according to paragraphs 8-9, characterised in that the mineral resources are copper deposits.
11. A method for selective arsenic removal from a variety of mineral resources, raw materials industry or soils, wherein the removal of arsenic is carried out by dissimilatory arsenate reduction using a bacterial strain comprising the plasmid pSheB, having the sequence shown in SEQ ID NO: 1 or its functional derivative, preferably the *Shewanella* sp. O23S strain deposited as KKP2045p, which method comprises the following steps:
- a) preparation of the mineral resources, wastes or soils and mixing with an appropriate culture medium enabling the cultivation of the strain,
 - b) addition of an inoculum of this strain and carrying out the culture under conditions enabling its growth and conduction of dissimilatory arsenate reduction.
12. The method according to paragraph 11, characterized in that, step b) is followed by step
- c) of selective removal of arsenic, released from the solutions obtained in step b).
13. The method according to paragraph 12, characterized in that, the removal of the released arsenic is carried out by flotation, preferably by selective precipitation of arsenites with sulfides.
14. The method according to paragraphs 11-13, characterized in that:
- step a) is carried out by: shredding and fractionation of the mineral resources, wastes and soils, preferably to the fraction of 125-250 µm.
15. The method according to paragraphs 11-14, characterized in that the culture medium is R1-R2 medium, supplemented with sodium lactate, yeast extract and Tuovinen salts.
16. The method according to paragraphs 11-15, characterized in that, the culture medium does not contain NO₃⁻ and Fe³⁺.
17. The method according to paragraphs 11-16, characterized in that in step b), the culture is carried out under anaerobic atmosphere conditions, and is carried out with flushing of the medium with a mixture of gases N₂:CO₂, preferably in a ratio 4:1.
- Having thus described in detail preferred embodiments of the present invention, it is to be understood that the invention defined by the above paragraphs is not to be limited to particular details set forth in the above description as many apparent variations thereof are possible without departing from the spirit or scope of the present invention.

SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 5

<210> SEQ ID NO 1
<211> LENGTH: 81591
<212> TYPE: DNA
<213> ORGANISM: Shewanella sp.

<400> SEQUENCE: 1

atgcaatttgc	ccaaatcg	attataaaa	tcgatttgc	agctagcgaa	tattgtgatt	60
tgcttggaaa	aaggaaagat	atttgaatcg	ttttttattt	caatcattgc	ggtaattccc	120
ctttttgc	ttgcttccaa	tacgaccaat	cagtcgttta	atcaggctaa	aaaacagctg	180
ttatcggtct	atcaagatca	gcgggaaacc	ctctattgc	gtgctgcatt	tgacgcaaaa	240
gggcaggtaa	tctctccg	aggatTTact	actaaaacac	attttagcg	cgctaaaaaa	300
atagaatggg	aacacgttgt	acccgccc	aattttggca	aagcatttat	tgagtggcgt	360
gatggccat	ggcaatgtgt	cgatagcaaa	ggttaagtgc	tcaaggccg	caagtgtgcc	420
gagaagatga	atgtttagt	tcgatacatg	caagcggata	tgcacaatct	atttccggcc	480
attggcgct	tgaatgc	gcfgcgt	tacaatttgc	cttgcatt	gtcggctaaa	540
tctgatttt	gtgcctgt	tatgcgcatt	gatggtagca	aagctcaacc	tcccgaaagac	600
gctagaggc	gcattgc	gtcctatcg	tatatggacc	aaagttatcc	taaatattcg	660
atgagtaagt	cgcagcgc	gcttatgagt	gcgtggata	agcaatatcc	ggtgaatatc	720
caagagtgc	agcgggctaa	aaaaattgc	gccattcag	tcaatgataa	tgagatcg	780
aaaagtgc	gtcagcaagc	caatattgg	taaacatacg	agaagctatg	aattatattg	840
ataataacac	cttgcagag	ttggatttgg	tttgggtgc	taagttaaa	gccagtaaga	900
atctcgatac	gctattaatc	caagtgc	gcgcgtg	taaaatagat	caagatc	960
cgtcaactca	acgtgagaag	tttgataatc	taacgtc	taatacgg	ttttgtctc	1020
gtgaaacg	agtaaaacag	ctcacggag	attggtta	aactggcgt	tccctcggt	1080
cgggctt	gttcgc	ccgcataac	gtgatcc	gatggcaagc	aagctgc	1140
tggctctac	tccctgc	ccaatcccta	acgcaatc	caagccccaa	acaaaagcaa	1200
acatcaactt	octaaagcac	tgtccgg	attggcg	aaagtgtgag	taagcgtgc	1260
cttctctatc	gttgc	tttccctgc	caatacc	cacgtc	ctgtca	1320
ttatcg	cattcacaa	cctgc	aggagg	tgcgc	gccc	1380
ctgtc	gtgaaatgag	tcaggaagtt	aggcgaagac	gagaacgata	tttaaacgac	1440
aggaaaaca	gcaatgacaa	actcaattca	agcgcaagc	actcacacta	aagccaccac	1500
tttac	agccaaaga	ctcaagcagg	cttcacaag	ccgatatt	ctcaaaatc	1560
accaactgc	gtcaacgc	tggggcatt	ggcttcaaca	gaaagccag	tgttactgc	1620
attgcagatt	aaccagg	tgttgc	gaagaacgc	cgtaaagaaa	atgcgtctaa	1680
ggctgc	gacagttgt	atgc	tttgc	ggcattttac	aaaatctgat	1740
tgtc	gagcc	atgaacgc	aaggcttata	cccagtttgc	ggcgggtggc	1800
acagctc	atttgc	aaagcgg	aaaataataa	gctgaaacc	aaaacgc	1860
gttgc	gaca	gccaaga	tggccaacta	tgccac	ctatcaatg	1920
cacacgtc	aagatgc	accgtc	atttgc	tttgc	tggtaataa	1980
ggcgct	attgc	tttgc	ggcgt	acagc	aat tcgtgc	2040
gcgc	catg	ttaagcatg	tcgccc	tgtgc	tgcataaa	2100

-continued

gtcgcttgat gttgtatga tattcacat tgccagtgc gaaaaacaag tcgaagtgtg	2160
ggagctggca ggcatagac gttacaacga aaaccaggtc cgcaatatgc ttaaaacgc	2220
cgtgtgaat gctgtactt atttggctca gtttggc caagaagaat acgaaaaagc	2280
gggggggtt gtgacccatg acttggc tagtgcgatc tatcttgatg ataaagcg	2340
gtcgaatca ttagcgaccg caaaatttga aatcgaaagca gcaaagttaa tcgcgcgtgg	2400
ctggaaatgg acacagataa agttatggc tgaatacgat gaacttgcag ggtttggc	2460
tctagactca aaagagggtg agtatgtatcc cgctgaaatg ggcgttgcgg ggtgtatgc	2520
cgtgttaaaa agttatggc agcctgtatc tatcttatatt gggtagtgc ataaggacga	2580
taaaaaggcg ctggcacaat taaaagccag tgtgcaatca gaccactca acgtgggtga	2640
cgtaaaaaaa attgaagaaa aggatactag cggttattcc gctgcgccta atgacgactt	2700
acggggcgacg cggttgcgat tcaacaaaca tgcgttaatg agtgcgccaa gcgttgcgc	2760
tgatacgttg cattttctg tctgtatgg cgcgttactt gattcacatt atggcagtcg	2820
tccactgcat atcagcgat acgacaccac ttgcacatcc aaaacggat cgctacgg	2880
taacaaagcg gtgcagttaa tcgagggtgt aaaagccagc ctaaatttag cgtgggtgg	2940
tctggccacg gtcgtgaac gcttaacgc ttttgcgc ctcgtatgaa aagagaaga	3000
aaagcaggtg gcttacgcca cgggttccat gttcgaggc tcaattgata atagtcataa	3060
agctgttgag gccgtgatata ccaegctcga tgtgaaatgg tctaaactact ggcgaccc	3120
tgcgaaacc ttcttaage gctgtatgg tctgtatgg attgtatgg ctcagccgt	3180
catgggtgag cagtggcat tacaagccgc tgcactcaag aaaaaggatt tagcaatca	3240
gggtgtatgc ttatgtacg gtgagcgaaa agggctatgat gatgcacaaa aagcctactt	3300
tgtatgtatgc atgcccacg gatccatgc cgcatactga tgcagatcc ccctgttaat	3360
tcaggtaaa tctgagcaag aattaactaa ataaatttag ttgtgtatgc gctgtgtat	3420
aatggttcaa cgggttacgc agagggttta caatgtatt tatggcggg atcgaagcc	3480
ctgtgtatgc aacgcaggcg tttgtatcg ttatccccgt gttgaaaaa ttaggtatgc	3540
gctgttttc ggccggcgat tgccaagaag aaatcttgc taaggctaaa gaagccattt	3600
tatataatggc tgaggaatgt atcaacatgc gtcattttatg cgatagtttgc aatgaaggat	3660
atcgccacta tcaggcattt catcccaatgc tgcaccaatgc gctggcctt gaggtgcac	3720
tagaggcgct taaagccaaa caaaaacgc tcaatatcac attgtcagaa tcgcagatttgc	3780
tacgtatcga tagcttgcg gcttttgc gcgagttcaaa agaccgttgc gacttcttgc	3840
ccaaggccgc cgataagtttta atgaaacacgc cagactcgtt aagatcgttgc tcaaaatgttgc	3900
gagattgttgc tggccaaatgc aaccgggttccat ccattttcag aggtgaaagc gcagcttgc	3960
gataatcccg aggttgcgc agcctatgcg caagcgttgc aagccgttgc accgttccac	4020
attatccatgc tgcaacacatgc tgagggcgat atcatgcataatgc tcaacaatatgc tctgtatgttgc	4080
actcgtgtatgc gtcatcatgt ttatccatgc ttgtatgc gtttggcgc catcaatgc	4140
cagtgtatgc gatccatgc ccctctggcg cgggtatgc tgggtatgc gggggcgatgc	4200
atcgatgcacg cggccatcaatgc tgccggcgat tgattgtatgc gtcgttgc gggggcgatgc	4260
gcttagcgatgc tttgtatgc ttgtatgc gggggcgatgc gatccatgc tttttccatgc tttttccatgc	4320
atgttccatgc gcaatgc tttgtatgc gggggcgatgc gatccatgc tttttccatgc tttttccatgc	4380
aagatgtatgc catgtatgc ttgtatgc gacaaggatgatgc attaggcatgc tctcaacatgc	4440

-continued

cgcatacgat	tgtaaagata	agtgtatctca	agtccgtcaa	acagggtgaa	gtgtttgagt	4500
ggtgttattga	tttatgaagaa	tttcaatggc	gcaaggggaga	tgactttcta	aggagtgc当地	4560
caggcgctcg	tttccttgg	gaaatggc	cgttgactga	taacacaaa	accgctgcta	4620
atcgcaaagt	ctttaagctg	ataaaatgaa	aggacaaaat	gatgtggcg	ttcttggata	4680
atgtactatt	ggtgatttgc	tggctgtca	ttgccatttc	tgctatttat	ttcgctcgag	4740
ccgcattcag	gtcggtttt	attggccgc	gatagtegc	attccctgtt	agttccccgc	4800
aaaacgc当地	acaagcggcg	gctgcgc当地	cctttccgt	ttatctgaca	atacccttg	4860
gtgtgttttgc	aaacgtgcct	cggcggtta	tctatattact	ctgagtgtct	attggctgct	4920
gtcgctat	gtcatctaaa	gacaactgat	gaacttggc	gttcccttcg	gtcgggctt	4980
tcgcttcgct	ccccgc当地	acgtgtatca	ctgatggcaa	gcaagcttc	cgtggctcta	5040
cggttctgct	tgccaatccc	taacgcaacc	accaaaccgc	caagcacagt	cctactcact	5100
gcccggcaag	tgtggctaag	ctggctgtt	cattcggcat	catggccgct	cctggctcta	5160
tcccttcgt	ttacgcttcc	gtcttatctg	tcgctcacc	caacccgc当地	gccagggtt	5220
cgcttcgccc	ctgc当地	ttgtcaatca	gaacgatggc	gagc当地	gacgtcaggc	5280
aaaaacgaag	atgatagacc	ccaaacggag	caagcagccc	atgacacact	caatcactac	5340
cccgcaagc	cacactaagc	ctttagcga	gcaagccagc	ttacaagcga	caagcaggct	5400
ctccaaagac	gatattcaac	agcaaaaaag	gggc当地	acgatcatg	ttttaactga	5460
aatgaatttc	tttacttct	gccc当地	acttgagggg	gtggctaatg	aggcatggta	5520
tccctattca	caagggttat	tggttgaagc	ggtc当地	tacttagtgc	cgccaaaat	5580
ggctcaaaat	gtcgatcaa	tcgatcaacc	tcgacgca	ggc当地	ttgattatcg	5640
cgtgcgtttt	aatgtgttta	gcactgacaa	aaaacgggtt	tcgggtcg	ttcctgtcg	5700
gattaagaat	tttggcaacg	atacgcaatg	cctctcagtc	gatgactaa	cgagggcg	5760
aagtctgaa	tatcgaaata	agcacgttcc	aatcatgtgg	tcttaccct	cggggatcac	5820
atcaacggtt	tatgtgtcg	tggcgctga	cgggggata	agcgattgct	atacacatca	5880
tgc当地	tttggatgg	aagggggtt	cattgagtttgc	gccactaaca	cgggggcat	5940
atgatatgt	caaaccctca	gtgacgttttgc	aagaaatgaa	cacctattta	caagggcaat	6000
ataaacatga	acgtttat	ggtc当地	ctgcaacggg	ttggctgc当地	gactatgca	6060
catcggttgc	cagtagccac	atggatgatt	tgaatgctaa	cgggatggg	tacattggcc	6120
atcacgaatc	cgc当地	cagggcgttcc	gttataccgt	ggacgagg	aaatattccc	6180
cccgctttt	tattgttggc	tatgactccg	aaaacgat	tgcccaaaaa	caatgctatt	6240
ttattgtatc	agcgatcaat	atcgaaaggg	ctcgatcaa	atgcttagag	agtgtatc	6300
getatccgat	agtgagaatt	aaggatctac	aaggcagtt	tgtgtaaaca	gcctgaa	6360
tttacagggg	aatcgatcg	ccc当地	gttaattc	atcatgaccc	acttcgttt	6420
tttcgtctta	agttcaaacc	gttc当地	ttaagtctcg	aaaatatct	ttttgcaac	6480
tatataatgt	tgcacttttgc	gccc当地	tggtataatg	atccc当地	tagcagggg	6540
cattatgacg	aaaacggata	aattattac	aaagctcaa	tcgtccaaa	cactgacttgc	6600
gaatgaat	gcggggcgt	taaaagtgc	cggttatcaa	caagtc当地	gtgtatgg	6660
gagggttaag	tttgataatg	gcattccgaa	ggagttatc	aacctgcata	aaccgc当地	6720
taaaaatgaa	ttaaaggctt	acgc当地	ccaaagtgaga	gacaaat	ccgaatgggg	6780
gaaaactatga	tgcaaata	aatgacat	aagggttatac	acggctctgt	agaaatcagc	6840

-continued

ccaggaggatattatctgtt tggcaggtgtt ctattcattt caccgttgat caactatgag 6900
gccgaaacag ccaaagggtt agagcaagcc ttcaagagg cgattaacgc ttatcttgct 6960
gactgtgctc aacaggacat tcaaccagag aagccttgta agggatcgct taatgtcaga 7020
ttaggccatg acttgcattt ggctgcgtct atcgcggcat ttcaaggatc gactagcaeg 7080
aataggttta tcaaaaaggc agttaaaaa agcgtgggtt tgttaggctgc aaatgttgt 7140
cgagcaatga tatcagccctg agccatgtga tctacttaaa ctgaacattt cccctgtta 7200
getccccgca aaacgcgaaa caagcgggg ctgcgcgc cttttccgtt tatctgacaa 7260
taccccttgg tggctttga aacgtgcctc ggcggttaat ctatttactc tgagtgtcta 7320
ttggctgtctg tcgcgtatg tcatactaaag acaactgtatg aacttggcg ttcccttcgg 7380
tcgggctttt cgcttcgctc cccgcgataa cgtgatccac ttagtggcaag caagcttgc 7440
gtggctctac gttctgttgc gccaatccct aacgcaacca ccaaaccgccc aagcacagtc 7500
ctactcaactg cgccgcaagt gtggctaagc tggtcgttca ttgcgcacatc atggccgcctc 7560
ctggtgctat cccttcgtt tacgcttccg ttttatctgt cgctcaccac aaccgcgtt 7620
ccaagggttgc gcttcggccgc tgccgcgcct tgcataatc aacgatggca agcggcagg 7680
acgtcaggcg aaaacgaaga ttagatgaccc caaacggagc aagcagccca tgacacactc 7740
aatcaactacc ccagcaagcc acactaagcc tttcagggagc caagccagct tacaagcgc 7800
aaggcaggctc tccaaagacg ataaattgaa aggctttaaa gatggaaac tacgtaaaaa 7860
cgcaaaaagt aaaattgcag tggatgttgc aatggcgaca aagatcgatc tcagaacccca 7920
aacaatgggg atggatgatt gtttaatga taatgccccg cagatcggtt atataattt 7980
cttttggagc tggatcgaga aggaataactc agagctagaa atttatgtaa ccgtgattga 8040
tggatgttata actaaagtgc gttttctga ttgtccctat cattttctg acgatatcg 8100
tttgacccctt gaagaagtc aagcgaaaaa gcccgcattt acctatgcac aagttaaaca 8160
ggcattgttta gacggcttatt tatgcccgtt gtcgaacgt tccaaagcccc ctgagccaaac 8220
accggccaaac gatgacaaca ctcgcaaatg gtttagttt ggcgattatc aagatcgct 8280
cgaaagcaag cgtgaccgct tagagagcgc cgcaggtaaa gcggccggcag actcaacaa 8340
attttatgag tcgtctagaa gccttgcattt ttgtattccg ttccggccagc caattttgtt 8400
tggacaccat agcgagaaac gcgcgagacg ccacgcacag aagatttttta atgatatggg 8460
taaatctgtt gcccgcagca aaaaagcggg ttattacgct gacagagcgg ccagcgttgg 8520
caactaatggc attgcacatcg attgtcccg aactatcgcc aaactaaaag aaaaactcgc 8580
aggcttggag cgttccacaag aaacgatgaa agcgatcaat aaagttatcc gctctaagca 8640
tatgacggac gcccgcacaaa tcgactacat gacgcacacg cataagctaa ccgaaaaaga 8700
agcggaaagaa ttactaaagg gggatttttgc cggccgggtt gggattttgtt gttactcaat 8760
cgcaaaataac agcgccacta ttgcacccgt aaggatcgatc attgaagatt tggaaaagtt 8820
acataatcaa gttgcatttgc gtgcacggg tggatgttgc gggattttttt gggcggttata 8880
tggatgttata ggcgttata aaataactt tgatgtatata ccaaggcagg cgctacgggtt 8940
aacaataaaa aaacatgcattt ttaaatggctt gcgatattctt aaagcttggg ttgcgtaaaat 9000
aacacctaataat gcccattata gcccacacaa attgtatggc aagttggcag aaaagtagca 9060
gtgtcaaaa tggtaatcc ttgtcaattt ttaatgggtt aagtggccgc tctgcgtca 9120
tctaattttac aactggttt aaatggagga ggttggggc ttaaaaattt cagaggttgc 9180

-continued

cgcgattgaa taaaataaga ttaatttcac cttgctatgt gtagatttttataataaaat 9240
tttggtagt aaataaacgg agaagtacgg cggggattgt gttgcacatca atcaatgtaa 9300
aatgccattt tttagtgtt taaagtttg cgattttac gaggcagcat tgacgattt 9360
aactcaccta gcaggatatt aagatcgaa tgatagaat aaaggattta acttcgctc 9420
atgttaattgc caattttaaa agccattcgc cgaatttagga gcaagttctt ttaatggtt 9480
agtttaagtgt taacggtaac gttcaacatg attcctatat agcaagttgg cttagaaatgt 9540
tggaaaggat aatcgcagca tctaaggcac atcattattt aatcgaatta acgttataaa 9600
taaaaaagcca tcttagttaa tgagttcagt ttaaataat agtaatggga tttttatgtt 9660
agaaacaata gaagaaaaag ccgcaagaga aaagagggtg ttagcgtggc ttggggctgg 9720
gtatgtctcg caaggcgtac ttgcgctaat cgaaaattca agcaatgaag aactggcagt 9780
attatttac cgtcacgcgt tggagattt ggggctcggt cctgatgagg ataggcaatc 9840
taatgatcac gcacttcgaa ataatgaacg ggtaatatca tcctatttt tcgcaggat 9900
taagattttg atcatcactg aggcagacag agcacgaacc accgcattat tgccaagcga 9960
atattaaagcc gccacgctt cagaaaacgc actcatttag tgcgttttt tggcttgaaa 10020
atcccccgat acctttgcgc cttcatacgg cagcagccta gcaaaaatcg aaaattccca 10080
aaaaatcccc aaaaatgaa aagtgcaaaa ccctggcgaa atagtgaagt tggaaaccttg 10140
agaataagac cacacccac gcaagaattt ataaatcaat tcttatggtt gtttgggtg 10200
taaatttttgcgtt ttacactaaa tcgttattaa ggtgttattaa taaaataata caatgtatt 10260
catttcgtt ttacactaaa tcgttattaa ggtgttattaa taaaataata caatgtatt 10320
aaaagccctt gtaccaatgg ttccaggccg attaatacag tattaatacc gtgtgttatt 10380
tttagtgtt aagtgtatgtt tataaaagaa ttacttattt actcattttg tgagtgtatt 10440
aaactttaat acactggta atacagttt atgcacaaacc gatacgttgc taatacattt 10500
ttaatacatt ggataatgtc actatggccaa gaagtaaagt tgtcaagtc ccctgcgtat 10560
cagatctcta tgccaaaggattt gttgttattt gagaagagaa ggggtttagaa aaagatgtcg 10620
ctggcgegccc agatttgatc ttgtttgcgc tgctgttattt tgacataaaa gatgtgtccg 10680
aggggttattt tactcgcgaa ctgttagagg tgattttgc ctatacattt aaggcacaac 10740
acacttagcag ttgggtctat taccagacat acaatggcc tcctgtatgt atacagaatg 10800
caggttattaa gcataaggac actatggaaa aggccaggaa aaaagtttag cagatttttag 10860
cgggcgaaaa taaagaggac taagtgtccct ctttatataat tcattacatg gtctttccc 10920
caatatccaa tttaggattt tcgtatgttgg gggagttttc ttgcgtcggtt gaacgggtggc 10980
tatccattgc taaaatgttga tagtgttcat cgtgtttttt tttattcgat gttacgttt 11040
tgctgtcatc tttcatcgat atattttttgcgttgcattt atcatgactc aatgattgaa 11100
tggcttttaga tagggatttttgcgttgcgttgcattt gatcgttgcgttgcgttgcattt 11160
gttttgcgttgcgttgcgttgcgttgcgttgcattt gatcgttgcgttgcgttgcattt 11220
catgtatattt ttcccttgcgttgcgttgcattt gatcgttgcgttgcattt gatcgttgcgttgcgttgcattt 11280
gtgaatttttgcgttgcgttgcattt gatcgttgcgttgcattt gatcgttgcgttgcattt 11340
tgcttgcattt attgatgttgcgttgcattt gatcgttgcgttgcattt gatcgttgcgttgcattt 11400
taagcgcgttgcgttgcattt gatcgttgcgttgcattt gatcgttgcgttgcattt 11460
tcgtcgatcgttgcgttgcattt gatcgttgcgttgcattt gatcgttgcgttgcattt 11520
tgctcaccgcgttgcgttgcattt gatcgttgcgttgcattt gatcgttgcgttgcattt 11580

-continued

tgggtatgtt cgcaattaat gcgggaaagg ttccacgttt ttctgacgaa taaacggcat 11640
 cgtgaaaacg taagtttca tgctgcttaa cgtcgatgcc ttgcttggtt aaatacgtct 11700
 gtgcgagcga gttttctata gggttgcct cgttaaagta ttgcttcgcg cgcgcttcta 11760
 attggcttg cttaaccgga atggcatgtg ttaattcata atgctgtggg tttttacca 11820
 gattatattt atcgggttcg ttcatcattt tatcggttt ataaagtgc tccttataac 11880
 tgcatttttc agcgaccata atcaggtaa ttaggggtcc ttatcgcgcg gttttccagt 11940
 cttaaaata gccacgatat tggccgttg tggtaatgt taggctccc tctttgcca 12000
 atgtcagggtt atcttgcgtct gatttatcgc gggtcggcgc acctaaaagc tgcgttgc 12060
 gegattcaact gtaggcccga agcttggcat ttagctcagt ggccatccgt tttagatcca 12120
 catatccatt cacatcgagg tactttgggt tttcctgcgg caagggttgc tcattaaata 12180
 gccccgtatg ttcaggtaa tacttatctc ttgtctcaag tgcagagact ttattggat 12240
 cgttattgac ccatgagagc atcattttt gccccgttatac ggtgatgatt tttacgtgct 12300
 tactggcactg ggtgatatcg atatacgcgc gccttacgtt agttaatgc ccttggccct 12360
 taatcgcagt aatgacattt tcatacggtt cccttgcgc catatcgccg gtgcgggtgt 12420
 atgcataatc ccaatgagca tctttcaact ctttgggtga gagcgttaat gtgtggtat 12480
 cctgtggattt gagctgcatt tgtttategtt caatgcggct aacggatag gttttgtgc 12540
 cctgcatttc acggtaataa tcgttttgc ggagcatgtt actgttgtc tctgegagag 12600
 gttgctact cttagccaa aggttgcgtga atttgcggcc atgatttttgggaaat 12660
 gtttttctc gccccgttatac atatccgcga gggtgagcat gttattgagc ttatcgac 12720
 tcataatttc aaaataactg tctttactgg tcgttagat taacccctgc tgataggca 12780
 acatggtcgc caattccgc ttactggcat gaacacccgcg taatctcggt acggtaatt 12840
 gttcggtact gaggttttt tggtgctgtt acccttgcgc gatttgtgg gtaatgtcgt 12900
 cgcggttctt gttgggtgtac gcaatgataa ggggttttgc acgcgttta ggtgttctgg 12960
 ccaagtattt cagtgcacc gtggggggca gggcttcggg cgccagttgc tggttttct 13020
 tcgggttatac ggttatacgac tgcaaagtgc aaatgatatt caaacttgc ttgtatgttag 13080
 cggtagctc ctttccggc gttaccgggtt cggtcgagg gggcggatgt tattggtaa 13140
 agggttaattt tgctgaattt ttatcaatag cgcttgcgc ttggcgatcg acaatattgt 13200
 gtactgcact cagtagcgtg ctgttctgtt ggccacaat gtcttcatg gtgcgttat 13260
 tgattgcgc tttgttaac gcgagttcaa aagggttacc cgcttctgg gacgttaact 13320
 gtttagtgta acccaagaac acggcgccgc cgccgttgc ctcaatcaaa tggtaaatt 13380
 tatccatttgc ggcgttactg gtcatcgagg attcatcgag caagaacacg gtgttagcgt 13440
 attgggtctgg ttctgcatact tggcttaaaa agttgcgttag gaggatgtgc acagttggc 13500
 ttctgcacgtt ctgttctttt gtttcgtttaa cggccggcgtg tgcggcgct agacctatga 13560
 attgtgtcgg tgcgttattt ggttgtgagt tgagaatggaa ttgcgtttgt ttaacaagat 13620
 cgatacccgatg ttcaaggatg gtcgttac ccgttccgc taacccctga acacccatcaa 13680
 atcgggtttt ggtgggtgtt ataagcaagg tcgcgtttt ttggccatcg gtcagggttt 13740
 gattttgcgtt gaggatgtgc cggccgtgtt caacgcgttgc aaggggcgca agttgatctt 13800
 tggccgttgc tagtctgtctt aaaatacgca ctgcgttac gatggcggttgc tgggtggtcc 13860
 acgggtgcc gtcgtgttat tcggcgctca ggagtgggg ccgcgcctta tccgtgtgt 13920

-continued

tcgtggccat gtgttcagg gtgttgcagg tttcatgttg ggtaatgctg gtgcctttt 13980
 cctcaaaggc gtacttgcgtc gcattaaacaa tcagttttt ttgtgagtaa ccggcttctt 14040
 ttcgtcaat gtgctggatg gcaaaatcca cggcattgtg aataagccct ttgtcttgc 14100
 gctgggcacc aatggcattc agtgcgtgcgg tcacatctt ggtgagcagg gccgtatgtt 14160
 ggcttattcac gaacttttctt ggcataatcg cggcttgcatt gatgcgcgaa atactggatg 14220
 gtttcaatc acctttttctt aatcggtcat caaatttgtt gccttgcgg tgaatataatc 14280
 aatgtgttttgc atgccttgc tcagcgtgtc atgttaaccac tcttttgcagg cgctataaggc 14340
 tttatatttgc acgtatgttgc gacttgcatt agggcttgcga ttgcttgcgc tggtggcgta 14400
 accatgatttgc agcggcgcat ggtgcaatgc actggcttta acgctgtgtt gtttgcgc 14460
 ttcgtctttgc agcgtgaccc attcttgcga cagggcggttgc atcgtaataaa ttgtttgc 14520
 caaaccactg tcttgcgttgc ttgtggccat gatgagtttgc tgcccggttgc cgatatgtaa 14580
 attgtccggatgc ttagcgatcg cgtactctgc ggctttaaag gctttgcgtc tggggtaat 14640
 ctccggacgc ttaccgtttt tgcgttatcg cgtgagggttgc ttgtttatcg ggtttacact 14700
 gtccaccgtc aattgtgtgc gttgattgcg gccaccgact ttttgcattt gggtgagttac 14760
 catgcccctttgc gtgtactggc cgacgacttc ccgttgcgtc gcccataaga acacgggatt 14820
 aagggttaggg atgttgcgttgc caaggcgcga tacgtcaccc tgcgttgcg 14880
 gatcgcgttgc ttaagctgtc tcacgtctt atgggtgcgc gcaatgactt ggggtgtttgc 14940
 ctctctggcgttgc aagtttttgc aatcaggatcc accctgttgc tgcgtttgc 15000
 ttcatgcgttgc cgcattggcgttgc tggttgcgttgc ttttgcatttgc 15060
 gttgccttttgc ttaagtaccc tcacgtcgatcc accggcttgc atgccttgc tggcattggc 15120
 gtgatttgcgttgc aaaatcaactt tactgttacc cgcttttgcgttgc tgccgcgcgaa gggcattaa 15180
 gtcgttaaggcc cgagtttat tggcggttgc aaccactaaa atatcgccgt tagtgtttgc 15240
 ggccgttttgc taaagctgc tgagtgtatgc ggtgtggcgttgc ggtttaaagtttgc 15300
 ccacttgcgttgc atagtcgttgc aatcccgttgc taccgttgc tgcgttgc gtttgcgttgc 15360
 catgtcagggttgc taaatgttgc gtaatgttgc accgttgcgttgc taccgttgc tgcgttgc 15420
 tggcgactgttgc tggcgactgttgc tgccaaagac gtttgcgttgc ttaacttgc ttttgcgttgc 15480
 gaacagtgcgttgc gagacttttgc ttttgcgttgc cgcatgttgc ctttgcgttgc gaaatcggttgc 15540
 ggtatcgacc acttggcgca tggcggttgc ttttgcgttgc atgttgcgttgc tgcgttgc 15600
 ttctgttttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc 15660
 tgcgttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc 15720
 gcttgcgttgc aattcaagcg tggcgatcc cactaatttgc gcttgcgttgc ttttgcgttgc 15780
 ggtgttacttgc gcttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc 15840
 taacgcatttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc 15900
 gccgttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc 15960
 ttgttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc 16020
 gcttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc 16080
 cgggacaccc tccacttcaa acaaatttattt acccagtgcgttgc ttttgcgttgc ttttgcgttgc 16140
 ttcaacactggttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc 16200
 aacccgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc 16260
 gcgcaatgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc 16320

-continued

cggatcttct tcacggctgg ttttgtggcg caccatcgca aacagcaaat ttcccgtgtt 16380
 ttc当地aaagag gtcgctttgg tatacggggtt ggtttgcgtc gcctgcgcgc catcttctc 16440
 gatatgcgt aacgcgaatt ttaccgcgtc atcgtgcgtc gtcattaagc gtttatcgcc 16500
 gcccacaggt gcaagtaaaac tgaccgattt aggccgcgaa aaggtaaagt caaaaaccgtt 16560
 gccccgggtt ctgttttggg ttttaagcgt ttggcgttc aggcgtaccgg ctaacacctc 16620
 ttggagttga tggcttcaa cgggcttacc tagcattccc tcttgcgtc ccagttgcc 16680
 aaaccattgg gtattggcgt cggcgctcg ctcttcaag taatagttgg tttcccttc 16740
 ttggcgtcc gctgtttttt ctaacgcac acgtcaggaaat ttggagttct tttcttcaact 16800
 gaggttagtag ctggccgcgc cgccagcgct cgacgcataat ggactgtatgg aaagcataat 16860
 gatattccct tagaggggacc tcgcgcgggg cacgagggtga actaatccgg taggtttatt 16920
 ctgagtcgag ttgggctttt atttcgtcag acgtgaagcc cagcgcttgc agttccttga 16980
 tggcttcaat tcgacgcgt acgttcatcc gttcattgtc atttttagag gctttacgtt 17040
 agggctttt cggccacaat tcgggcactc gtactgtggt gttgaatgag gtcatcggtt 17100
 tagctcctt agtcgagcga tatttcttca tggactgagg ctatttgcgt ctcttcaagt 17160
 cgctgtgcattt octcaatgtt ctggcatttgc gctttttgtt tttcagggc tttttcttc 17220
 tctgcattttt tgagcgctaa ctctgtcggt ttggcctttt tgcgcatttgc gaggtttgg 17280
 gogacccat cggtttttc agcaagtttg gtcatgcgtc gogctccgc gctgegtca 17340
 tcattttttt catcccttc atacatgggt tcatgaatgc ttggaggtt ggtttatcg 17400
 gccttacgat gatgttgcgc cgcgcgttgc tgggtatgt taaggcttatt ctctaattt 17460
 ttcaacttctt cagatgggtt gtaatccgc ttaataaaacg gggcgcacag tgagcgattt 17520
 tggtaaaga ctaaatccac tcttaccacg ggataattgc caggcgtgcg cagccaaat 17580
 tgtaaatctt caagctgcattt gatttcagag ggcgtgacca caggcgggtt aattgttga 17640
 tggcccagtg atacgcattt cctgatttca tttagcgcgtt agctgtatgtt ctctttagag 17700
 acatcgactt cttgttcacc taaatcgat gacgagattt tggccatttc gtttagaggtt 17760
 gagcggaaaa atagecgagt gttgagcaag tcgaacattt caagcgccgc atttcgcca 17820
 tagactttttt taagctgcgc aaaagactga atacctatga cataacagcc gccgaacttg 17880
 cgcaacttctg cgatttttc ggcaagttca ggtaacttgc gcaagttgg tgcttcattcc 17940
 atcaacaccc aaattctgcg gtccggatct tgcgtttgtt cttaaatggc ggtcgaggca 18000
 atggatagcc acattgagat cagtgccgc aatgaagcat gttgctgtgc gttactggac 18060
 agaaatagaa agcccttttgc ggcattttttt ttcacccactt cacgaataga aaatttgcgg 18120
 cgaagaagttt tccctgttc gtcctttca tccagtcggc caaggaatcg taatgatttgc 18180
 atgttagggccg cgaggaccga tttaatcgat atcgcgcgtt ttggatattt atctgatacc 18240
 atgtgaggccg attcggtttttt ttttaagtat tcaacttaagg actcaagtcc gctggcgtc 18300
 atgagttcaa gcagtcgttc agtggtttgg ttcttttagt gttcatcttgc catcttaat 18360
 gcaatggatgat aaaaaatggt acgtgcgtat tgcacccaaat gttgatcgcc ctcgcgtgt 18420
 tggggaaatga ggcgcgtcgc catgttctca aaatctgagg gttcatgtgc atcacaccac 18480
 acatcccaat aggcacacgcg ttcatcaaac gggtttagt gtcacgtccgt tgccggatata 18540
 taaaattttac tgacaaaggtt acagcccttg tcatagataa tggctttgtc gccccgttca 18600
 cgtatccacg taaaaactt acgcagcgcc acggactttc ctgcgcggcgtt ggtgcattca 18660

-continued

agtaacaagt gttgtacttc aaagcggttct ttgaaaatac cgataccatc gagtttgaaa 18720
teggactggc gcttgcctt ttaagctgt tttgctaatt ccttcggcgt ggcttttgc 18780
atcccccgaa taaagtgttc ctgcgttgc tcatcgccct ttttcttgcgaa ataccacatg 18840
gcaactgaca ggaaaacgat gccaatcaca cagctaataa ggaggtaaat ttgtgtccat 18900
aaccaaaagc ggctgttggc cgcaatcagt atcgggtttt ctaattgaga ccctaacgta 18960
ctggataga cagtgccctt ccattcagtc tgcatatacat gactttctgg tctaaacacg 19020
ctaaccgttt ggtttatcca ataataataa gcccgtta aggtgtctt atcaaggaag 19080
agatagcata gcagecgctgt taagatgago acgcacaaac acgtccagcg gacaagcagc 19140
gcgttaatct gcagaaacat ccgcacattt tgaataaaaa tctggccacc acgggtgttag 19200
ttactgcctt tagtacgtt ggggtttaag gtcattagtg aggctccctt tctaaagagat 19260
ttcaataaaa ggagtttaatgt tttccatcat ggcggatgtt gtcatttagc gctttttgt 19320
tagtatcttg taagtttgtt ttacttatgc tttgaacttt agctataataa ggcttaatag 19380
taatatcctg aatttgtaat ttcccagagt tgatgttgc aatttagcttgc tacaagcaa 19440
taattttttt atcttgcata tcaaattttt tttttaaata cgggtggcata atttgctcg 19500
caaaattttc agattttgat ttaaccaaata cacttcgtt aagttttt agtaaatcat 19560
ttcgttaattt taaattttcg atagtttctc cagcaaaaaat cttgctgcaat ttatctggat 19620
acagtttaca aatgaaatct ataaaatctt gagtgaagtc atagtttgcg atttttggc 19680
cgcttttgcg catacttgcg atgcttagat agtaatctgc taagttttctc tgctttttgtt 19740
gacctgtttt gtacttgcgat gtgttaataa tacctacaag aaaattttt tcttgcgcg 19800
atagccgtgc tgcttgagat actgaactta ataagcaataa taataatagg catgttttta 19860
tgatattcat taggttatct ctttttgatt gaaggctatc gctattatca catttttgggt 19920
cctgtattaa agggacttgc ttgacgcact gtttctgcata ttttttcata gtcacttcg 19980
ttgtctggct gtatttaggt gtgegcgcgt tttgcgtcg cacgggcgtt ttccatatac 20040
atgctcggttgc ctgcgttgc atggccttgg atgcgcctaa attgcgcgtt ttggtagagg 20160
ctgtcatttgc acggcgccagg ggcataactgt ggtgattgtt gtttatttgcgaa atcaatatcg 20220
gccccatgc gctccgtggc ttgcacaatgc tcatcgccgtt tttgcgttgcgat cgtgcgttt 20280
ttgttattgtc cagcaatggc tttcacttca gggctttgc taaattcgcgat cgcaagtttgc 20340
gttcgttgcgtt gctgcatttgc tggcggttgc cccctgtatca cccgtatgaa atcacattgc 20400
tcatcttttgc atcttccaaac gaattccaca aagtcttgcgat taaaatcgatc actaataactg 20460
atagatccac ttttcacgcg attagccgcgat tgggtgttgcgat tttgggttgcgat agtggactgg 20520
gattggagat ttttggcgat ggttgcgttgc tgggttgcgat cttgttaggaa ggtttttgtt 20580
tcacccgttgc ttttacttgc ggttgcgttgc gcttgcgttgcgat cgtgttataa ggatttgacg 20640
gtgttaacgt tagcgctgaa ctggtcaagg tagcttgcgat gtagactgtc acgttgcgtt 20700
tcactggatc gggcggtgtgc gacatcactg gttgagccgtt tagtgcgttgcgat attggcgat 20760
acgtgcgcac tagaaccatc aatccactc ccaggtgtgcgat cggatccac ccctgttgc 20820
gtcccgacat acattgataa taaggacttgc gttggcttgcgat cttctgttgcgat gttgaagtc 20880
tttagccaaatg acttaaccgttactacgcata tctgcgttgcgat ctttgcgttgcgat ggttgcgtt 20940
tcgttgcgttgcgat tcttgcgttgc gggcggttgc gttggcttgcgat cttctgttgcgat gttgaagtc 21000
agtgcggcat tcttgcgttgc gggcggttgc gttggcttgcgat cttctgttgcgat aattttgcgtt 21060

-continued

gcaactgatcg cggttgcattc agattgcgt aaggagtttgc agaccacatc gttactgctg 21120
 atattcaactg gcaagtcacg gcttaaagca tttggccat taatgacttg tctggccatca 21180
 tcaaagggtgg tggttcata gccagttgc tcttgcttgg tggtgcccatt aatacggtca 21240
 acgccagttgg tgtcgtttt attggccgttgc gtgttgcgttgc agctgttagt atccacccatga 21300
 gcggttaccca aattgatatac accacttgcc gcactcgca aggtgcgcgc gttaacgcgt 21360
 ctcatcatgc ccgcgaactg atgcgacata ctgctcattt ccgcgcgcgc gccttcata 21420
 atgagcggga caattaatgt cggtacgttgc atcattaaaaa agccggcat ggcgcata 21480
 cgggagtgc tttcttgc tttcttgc tttcttgc ttgcttaacg tgataccgc gtaaatgtt 21540
 gtgatttcgg tggcgtaag ggataagcgc gtggcatgtt gtaagttat gagggcaaa 21600
 atcacccggcc atgtgccttgc gtagatataag ctttgcattt agttgcgcgc cacttgc 21660
 gtgagtgagg ggatcatggc aatcgccacc actaaaaagg cgacagagct aaacagcaaa 21720
 aacaaaatcg attgcgcat gggcaaaaat tcccgccctt gcagggccat tcccgccac 21780
 atcgatgtgg tttgcattttt gtttgcgttgc taggcgttgc tgaatgcggc ggcgggttt 21840
 ttgctcaccg ctgcatttgc tgcattttgc tttcgccaggcg ctgttgcatttgc catgtttgc 21900
 agcgtgatggt ggcgcgcgtt ttgactgttgc ttggcgttgc tttgataacc actgtgtatg 21960
 ctgggtattca aaaagcttttgc ttggcgttgc gcttttcac cataaatccatc actgcgcgt 22020
 aacttgcatttgc ttttgcgttgc gtctgcatttgc aatagctgttgc tgaggcgccg aagggttct 22080
 ttgcaggtagt gatagtcatttgc aggaccatttgc tgaatgcgcgc gtagtggcga aggggttgc 22140
 gtcgcgttgc aatcccaat atccggcgca ctggcttgc ttcgcatttgc gtatttatttgc 22200
 ttgagtagca tatctcccgatc aatacaacttgc ctgtatgttgc tgcattttgc gccttgcatttgc 22260
 tgggtttttt ctatctcgatc ttgtctggat aacttgcatttgc gtcgcgttgc aaacatcatg 22320
 cccgttttac tgcattttgc atcatccgttgc acgtggggaaat tatcttccac cgtttgggttgc 22380
 gtggcataca tataatgc tggccatgttgc gccggaaatgc ccactatcgcc gggcatttgc 22440
 tccacgcgttgc agttaccgttgc cggattgggtt aaatcggttgc ttgctgttgc ggcgggtttgc 22500
 ttcatggca gcaatggcac ggcggaaatgc accgcgcacc atttgcatttgc cggttatgg 22560
 tcacgcgttgc tggggatggc tggggatggc tggggatggc tggggatggc tggggatggc 22620
 atcatgttgc tggggatggc tggggatggc tggggatggc tggggatggc tggggatggc 22680
 gtcacgttgc tggggatggc tggggatggc tggggatggc tggggatggc tggggatggc 22740
 tcaagttggc ctttagtttgc gtcgttgc tggggatggc tggggatggc tggggatggc 22800
 ttgcataat gagctgggttgc tggggatggc tggggatggc tggggatggc tggggatggc 22860
 cttaaaaatcg gtttgcgttgc ataatgttgc ttcgttgc tggggatggc tggggatggc 22920
 tggcgttgc tggggatggc tggggatggc tggggatggc tggggatggc tggggatggc 22980
 taaaactgcac gtcgttgc tggggatggc tggggatggc tggggatggc tggggatggc 23040
 attcggtacg taaaagggttgc atgtgggc tggggatggc tggggatggc tggggatggc 23100
 tatctgttag tggctgggttgc ttttgcatttgc ttcgttgc tggggatggc tggggatggc 23160
 cggcggttttgc tggggatggc tggggatggc tggggatggc tggggatggc tggggatggc 23220
 atgggtcaat ggcttgcatttgc ttcgttgc tggggatggc tggggatggc tggggatggc 23280
 tattaaatcg gtttgcatttgc ttcgttgc tggggatggc tggggatggc tggggatggc 23340
 cgtcatagat atacgttgc gaaatcgaca tggggatggc tggggatggc tggggatggc 23400

-continued

caaggaaggt gtttttaaaa atggcgcccc acacaatatt ggtgttttg cgacccatgt 23460
 cttttagtgc gggatcttc tttcgcttat cgagctgatt gttagcttga ccgccccgc 23520
 cgcattctg ctggccagat acccaatccg caaacgcgtt attttgtgtg cccatagtgg 23580
 cacaatgtt ttttgactg ccaacacctg caaaccccg cagtgcagat acgcccgtt 23640
 gggcactctc gcaagaattt atcgactgtat tgagatactt atcagcaatg gccgttagcg 23700
 tgccttaat ctgtttgatt tgccgtgecc acgtctgcag ggctaaatcg actgegaacg 23760
 gtgcagcggtt ggcgataata gccttgcett gttgcacaag ggcgtcagag ttaatgtggc 23820
 tgaatccccca cataaacata tcaataccgc cacagcctgc ggcaatgcgtg ggcatttgca 23880
 cactcacgag ctgcgcgtcg cgaatggggc tacgcacaaa gagattgcgg ccgcgtgttagt 23940
 agttggcgct ttggcccttg aatgagcccg gattggtcac gttagtgcac tagcctaggc 24000
 cattgaaaaa atcgtttagt gaacctgaaa cccccgcgtt ataggcgtgt gttgtcagca 24060
 aacatgaaat agtgacgccc agtagtgcgtc ggcgtcgtgc taagcgtttc atgatttacc 24120
 tccggatattt cgcgttatgc tgccgcttat gttgcgtta tgacgcgggt ttagtgtgcg 24180
 taggcgcacgc aagcaaacgg cacttaacaa tgcacgcact gggacgcgtt cgaatttata 24240
 aatggcataag gccagtaaaa tattgatgt taacatcgatg tccatgatgt tgctcctaacc 24300
 gcatgcgtgc aatgacttgtt gggtcattgtt gggattgtt gtaactttgt tgaagctgcg 24360
 agggcgtcactc atcaccgcacg gttaaacggc caaaacttacg ggtattgcacg ttgattaaaa 24420
 atgtggcagg cactgtcactg ctgcgcggat tctcaaagaa ggtggcggca atgtctggcg 24480
 tggaaaggat cggacttca aagcctgcta tgccttggtt gtcgagcgaa aaggcatagg 24540
 tataggtgcc agtttgcgcc gcaagctgtt tcagttcgg ggcgaacttgg tggcagtaag 24600
 ggcaatcaga gcgaaagaaa aacactaactg cgtaatcatt cggattggcc gcaactggcc 24660
 cccccccat aggccgtcgcc tttagcgctta aagagaatag aagagtctt aacattactg 24720
 ttagatgcg tttcatgggtt atttccgtt taaaagttagt gggcgaatc ggtggccacg 24780
 ttccaggaagc ggccaagtaa atcgcttgcgagataaactg catabcgag cggtttcatt 24840
 tccccctgc caggattaac taacacgtt gccggactaa acggggacgtt tatgtgtcgg 24900
 tegttttggc ggttctgcac gatggctggc agtgctacac catcgacgtt gacgcgtt 24960
 agctcaatgc catgctcagt agcgaactgc tgcacactgg ggcgtacgc aatatcaatt 25020
 tcttcttggc ctttgcgtt aaagaacacgc cccatcccg catggcaag gctgcgtact 25080
 ggcgcgtatct tcttgcgtt ctctaatgtt aaatagggtt ttcttgcgc ttgttccgtc 25140
 ggtctatctt tgggtgtt aagtcgtt cagtcgtt ggttgcgtt ttatcttgcgaaatc 25200
 cccaaactctg acgattttt gctgacgtac tgatgttgcgtt gcatggccct ttccacttgc 25260
 ttgcgtgtt ggttgcgtt cgcgttgcgtt ttgcgtgtt ggttgcgtt gcatggccct 25320
 caatccatct gctctgtggg cgttgcgtt ttgcgtgtt ggttgcgtt gcatggccct 25380
 tggccctttt taggtgcgtt cttttgggg ggcgtcggtt ctttgcgtt ggttgcgtt 25440
 tgcgtatccg tggcggtcgcc gcccggccaa aacgttgcgtt gcatggccct 25500
 ctgcgttgcgtt tgaccaccgc atccccctttt tttgtggaca accacgcggccaa cggcgccgtgg 25560
 ctgcgttgcgtt gggccgttgcgtt ggttgcgtt ggttgcgtt gcatggccct 25620
 atcatggttt ggttgcgtt ggttgcgtt ggttgcgtt ggttgcgtt gcatggccct 25680
 gcaatgcgtt ggcgtcgat atgttgcgtt aaaaatcaga aagtcgtt ggttgcgtt gcatggccct 25740
 caatgcgtt ggttgcgtt ggttgcgtt ggttgcgtt ggttgcgtt gcatggccct 25800

-continued

agttttgcc caattgcgcg cggccctcaa cctgcacaat gcgcgcgagc ttgctatcaa 25860
 acgcgcata ggcttttgc ttgcggatac acacgccaag cactttctca gcgcaatatt 25920
 ggccaagtga tacggtgagc ttgtcttctt tggcggtacc cagcgtttt tcttcttcac 25980
 tgcaggacgt taagccgatg ccgttacccc aaccgctgtc ctgacaacaa ttgctgatcc 26040
 ccagcgctt atcgctgcac tccatgatgt tgcccttgaa aatgatggcg ctgtttcgt 26100
 ttagtgttgc ttgtcttgc acatcacgg ctgcgtact gacggggcc agtctgccc 26160
 cagcatgggt aaagtcatgt ctgcgagtgg cagcaggatc gtaatagtc ccaactgagcg 26220
 caaagctgtc agcacccgcag ataaggcccg tatcgccgc ttgctgttta tcacataggc 26280
 gttgttgtt atcttaatg caaacgcgcg caagttgggt gtcacaactg gtcgattgaa 26340
 tgctgcaatc tgcgatggcg tcgcggatgt tgccaaaggc acattgatag gtggttgtt 26400
 cttccagca atccaatgtc acggctatcc cattaatcat ggcgcgttct ttaccttcta 26460
 cgcatggata ttggacttgtt tgccacttag cagaacacgc accacagctg ttgttgtaac 26520
 cgacttcaac atgcgtttca gtgggtatgaa tagtgcgttgc taaattctgt aactcaattt 26580
 taaaatgtt aatcccaatattt gcacttatata gaaccgcata tggggatgct ttaatatcat 26640
 aattttctgt ggttaatgaa ctgtttacac tcattgcgcg aggtgttgc gcaacacatg 26700
 taccaggggc tctctctgtt agaacatgtc ctataaacgtt accatttaat gcaattgata 26760
 attcggcatac aatggctataa gaacgataact cacctttaaa ttgaacacgg acgatctaa 26820
 cattagggtt tggtaatttata atgtcaatag taaacgggtt tgccggatgtt aagctgtatg 26880
 aatgcgcac ttccttaacc gatgtgttt tgacaaagggtt tgtaggttag cattgcaccg 26940
 gattattgtt gggcaacgtg caagaacgtg gcagagaatc gacttcacaa acgagccctt 27000
 tttcacagtc atgatattta gatgaaatgc cgtggctgat attgtaaagca tcattttgat 27060
 agccaaacggc ggtatgttgc gccggatgtt tttcatcaat cgtccggccg ccattattaa 27120
 agccactcgat cacggcttgg gcttggccat ccgtgctcat gccttgcatt ttggactgt 27180
 ccatggcgct cgggttggtagtgcgcg tttcgctgg gttagegact tcttgctgac 27240
 atgtcgcatc ttgcataagg ctgttaacgtt caaagccggt gtaaccggct ttgttctgt 27300
 tctgcttgcgt actgtctttgc gcccactgtca cattatcgat gtaatcttt tcgctgttgg 27360
 ccggccataac aggcaaggcg ggcgtcaatg tgaggatggc actgaggatt aaggcatat 27420
 tcaactgagt catttcgccc ccttgcataatgtt gtcacatcg gaatgctgccc 27480
 gtgttgtgca agtaagttca ggcgtcata caacgaaacg ttgcccgtaaa tcacgtcaaa 27540
 gtcatcggtt ctgcagggtt gtttgcctg acatttgcgtt ggtcttacag caacaaaggc 27600
 gggtaacttgg gtaatgcccattgcgttca ccactggggataatgc gacgcgttgc 27660
 tatggtttttgc cccctgttgcgttgcgttcaatgc gtcgttgcgttgcgttgcgttgc 27720
 ttggggcgtt acaccgcgcac caatgagtggc cacctgtaaatgcgttgcgttgcgttgc 27780
 ctgcgtcaaa ctgggtcgacgcgatccaa cgcacacaat accatcacac cattagcatc 27840
 ggcgtctggggtttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc 27900
 gctttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc 27960
 gggtaatgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc 28020
 ttggggcgtt acctgtgaaa ttgatccaaatgcgttgcgttgcgttgcgttgcgttgcgttgc 28080
 tttaaatgaa tggtaatgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc 28140

US 9,328,397 B2

49

50

-continued

tgttggtaat ttggctcaa acttttcat cgaaaatccc ttccttatt ggtcacaaaa 28200
 acacgcatt gcgtttttg aagttgataa accaaagggt gtcacccgac acagggttgt 28260
 cgtggcact ttccagatc accgttgctg tggtgtaagg gtgcgeccat agcgeatcg 28320
 gtatggtgg tgcatttga tagegatgc gggatttagg catgaggggc atcggcggtt 28380
 gatagcaaat cgcgcacatct tcaccgtggg attccatatac gatgccttgg cgatgcatt 28440
 ttagtttcg ggcgtctaaa atcaacgtac cgcctgaat tggcgtgtcg cggtaatgg 28500
 tggtgcagt cagggggtaa gcactgcctt gtgagccat acaccaaaac agaaaatcaa 28560
 aaggcaatat gctgttagtt gaggtggcca cggcttcggg tgcacaggcg agttgtgaca 28620
 cgatattgcc aaataaaatg gttcgggat tgaggataaa tgacaactcg tcatcatccc 28680
 acagtgggtc aatctctgtt aaataggca tgcggaaatt gtccgtggcc atacaagctg 28740
 cgcttgcatt aatttcgcgc caatagatca cagggtattt gtaccagtgc ccatgataaa 28800
 acgcgcacatc actggcattcg ttatcgcgt tagtgacgcg gccgcccatt tgggggaaat 28860
 tactcatgtc gagcttcatc cccatgttga ccatgcgttgc agggacgcgc gtcacgttca 28920
 ctaaggatc aggctcccaa tagccatat tgaggccat ttgcataaaaa attggtggcg 28980
 gtttggaca aaatgaaatg ggtgatacgg ggttgcgcgt atcgggttagg ctcccaggaa 29040
 taataggcac actgccaatg gtcatggga agatacgcg ccaacaaatg tcggtcatcg 29100
 ggtaacgaa cgtgccccgtt caggccgtt ggcgcgttgc ttgtgcggga atgactccca 29160
 ttaagaatgc taagagcagc gcagggtttaa tgcgtcgcat ggttgcgttc ctccatgaat 29220
 acggggaaat ttgacacatc tatttcgcgc acttgccaca tcaccccccgc ttgcataatg 29280
 acactcggca cgcgcattaaatg tgcgtatgcg cggctgagggt ttccttgcgt gtcaaagtaa 29340
 acgcgcgcatttcaattggcc atgcgtgttgc ttgggttcgc cggccgtttaaattccatgg 29400
 atgggtgtgt gcttgcatttgc ggttgcgtca aactcttgcg cccataacac ttgcgtgcaca 29460
 tcatccccat caaaaaacac cagcggtttaa ctgactcaa acttgcgcac gttggccctt 29520
 tgcaatgtcg gccatgttgc ggtgtcaat gggttacgc ggttccctt gggataaattc 29580
 aggttgcatt tagcatcatt gatactttgc gccagcgtttaa aagtgggatc gactaaaaac 29640
 acttctgggt ttttagttgt gataagtccc actggacgag ggcgcattaa accggcttg 29700
 acattctctt ggaattgcgc ggcgcataactg tctagctcac ccgttttttc taagggttgc 29760
 agacgggtttt caatccactg cagcatgtca atttcagcta tggggAACAC ttggccacg 29820
 gtgcgcgttgc acgcgcgcattt ggtggtaag agtgggcata ggcgcgtcat cagaagcgcg 29880
 aattttattca tttagccgttgc acctcggtcg cttggcgcgc gttcgccttgc ttttgcgtt 29940
 gaagggtttt cagtagtgac tgcgtaaattt cttgggtgcatttgc gactgcgcgc ccactgacga 30000
 cggccgggct gactaacacc acgacatggt tatcaatggc gtattgtgc accacatcat 30060
 cgagtgtttt ggttgcatttgc gttgatccattt tttcgcgttgc ttcgcgttgc gactgcgttt 30120
 gaccgataga ttggtaaag ctgcgcacgg tttcgtaat gtcataactcg acaacactgg 30180
 gttcgcgcgttgc caccacagggtt gttatcatga cggccagcgtt caccacccat aaaattgtt 30240
 gtatgtatgtt ttttccattt ttatgcggcc ttttgcgttgc ttcgcgttgc ggttgcgttt 30300
 gcatcaaaggc gggccatatac atcgggatag aatgttggg ctgtcgttgc aatggcatcg 30360
 agtaacggca tgccgcgtttt aatcaagttt tgcgtatccattt ctttgcgttgc ggttgcgttt 30420
 gagagcatgg cgcgcgcacca tggatcggtca aaaatccggt gaaaggacac tagccgcgg 30480
 gtttgcgttgc tgcgtatccattt aatcaagttt tgcgtatccattt ctttgcgttgc ggttgcgttt 30540

-continued

tgttgctcga agggcgagaa gtggtcgaga tgcttgcata agaaaatcccc aaaggccatca 30600
ccttggcgca atgtgatatg aatatctgaa ctgttcagtg cgcgttcggc ctggcggttgc 30660
gcaaagaagt cgttcatgcc ttgcgttact gtggcaaaagg tcggccgaa cttacgcact 30720
gtgcgataac ctgcgttgc gaaatcccta cttgcgcatt tcgcggcgct catgagtgc 30780
catgcttctt caatcacaca gatttcgga atagaacgtt ggcccgatata gtacattgt 30840
tggctgatcg tcaccatcaa cgcgaaaatc accggacgtt gtagctgc tttaaatcca 30900
tcaagttcca gcgtggtgat atcgatatttggatcgagca tcgagggtt attgaaaatt 30960
gcccgcattgaa tgccctgtga gcagaacttgc tttaaattgcg cggcaatatac actgatacgg 31020
gtatcgatcat tttttgtgc ggcaatgtca tagagtgc tttgcacatc gtcaatcagc 31080
gtgtcggttac cgcatatggatccatgcacgc aaaatagcat cgccgagtaa tgccgtctga 31140
aatcccggtga gttttcatc aggtgatgcg atggtggega taagcgggtt gatgtttccc 31200
aacacttcgg ttagtggatc aatggcttgc cttcatcatcgaggattgg gccaaaagta 31260
cccgcatccc gcacgatcaa tactttgcct aagtgggtaa acgggtttag gaaaatctgc 31320
gaactgtcga gatagacgccc gcctaattgtt tgagtgttgc ttttataact ctgacccttgc 31380
tctaaaatcc atgccttgc cccttggca aaaaatcgatt ttactaaaggc ttgcattgc 31440
aatgacttcc ctgccccctga gcccggcgtatgcattgt tttttttact ctgcccgcag 31500
ttaaacggat cgaaatagct aatttgcatttgc cgcattgttgc aagagatgc ccctgtgtat 31560
aagcgtttgc agtctgcaac aatcggttgc aatttcacca agtttgcatttgc tttcatcatgc 31620
aaacacaggc cccgttttgc gctgtcttc ataaaggctt cgctcattgc gaaacgggttgc 31680
gtggacagggttgc acatcgatgc ttgcgttgc tttttttact ctgacccttgc 31740
cggaagggtcg caatcgctt tgagggtgc tggcgttgc tttttttact ctgacccttgc 31800
gtgacattca atgtcatgtt ggtgttttgc aaggcgttgc agtgcattttgc tttttttact 31860
tccttgcgtt cagccagtttc atcccgccgc ctggggatttgc acaaccgcatttgc 31920
accgttttgc tttagtgcatttgc atcgatgc ttgcgttgc tggcgttgc tttttttact 31980
ttaataataaa aagacacgc gacacgatataa ggacatttgc cactgttcat ggtgtacttgc 32040
agtgtgcatttgc tttttttact tttttttact tttttttact tttttttact 32100
ccaaaggatgc tttttttact tttttttact tttttttact tttttttact 32160
ttgaccgaat caccgatataat gatgcatttgc ctgttcggc tggcgttgc tttttttact 32220
ggttgataact cgtttatatttgc aacgcgttgc acgcgttgc tttttttact tttttttact 32280
tgctcgccgc aatgtgcatttgc caaatcgatgc ggcgttgc tttttttact tttttttact 32340
tgcgccttgc tttttttact tttttttact tttttttact tttttttact tttttttact 32400
gttgcatttgc tttttttact tttttttact tttttttact tttttttact tttttttact 32460
ttggtgccaa agcctttggc tttttttact tttttttact tttttttact tttttttact 32520
ttatcgccaa tggcatttgcg accacttgcg tttttttact tttttttact tttttttact 32580
acttgggttgc tttttttact tttttttact tttttttact tttttttact tttttttact 32640
cacaccaagg cgttcatgtt tttttttact tttttttact tttttttact tttttttact 32700
gcaaaggccaa aaccaaggcgtt tttttttact tttttttact tttttttact tttttttact 32760
cgataggccaa attcatgttgc cagatgggttgc tttttttact tttttttact tttttttact 32820
ccgagatgttgc tttttttact tttttttact tttttttact tttttttact tttttttact 32880

-continued

ctacccgtcc aacggctt atcgaccacg atgttagacat cgctggatc aacgtaatca	32940
ccttgctcat tgagatacgg gaagattgtt actttttgg a ctgactcaga ggtgegcagc	33000
ggtacaccgt ggccgtcgcg gcgtggcaat gggataaaatg cttgttgtg gtcgcgagc	33060
ttgggaatgt cagttgtct ttggccgtt aaggccgtca ccgcagggtt ttgtgttgt	33120
ggctgattga acacaccact gtccgttaagt tgccctactt ccatcatatt ggtacagcca	33180
ttgtatgcgc ccacccatc gcagaaatag tcaccatcca gccctgctgc acagecactt	33240
aatagccatg aagccagtgc gataataccg gtgttacgta atgcatttt gtcatgatt	33300
gggttcaaaa tgcactcct taaatgaatt aattcgcacg gcacaaaggct gagttacag	33360
ggggggcggtt accttgggtt ttatattgtc gctgtgggtt aggctgttt gcgccttggg	33420
cctgcggctt aatattccct ggcacatgtt ctaacagtgg gttgctgata gtgttggagga	33480
tttgcgtgtt ggcactgaca ggattttggg cttgttgcgc tagcgtttt tcataatttt	33540
gaatgccaac cggatcgagt gggaaagcctt tcaaaaaacac gatgttgacg gtattgccc	33600
gatttagctc aataatcgcc tggttattgtc cagctaattt gatgttagtaa tcagcaagct	33660
tgcataccac gcttgaactg gcactgcccgg ctaagtttag cagtgcatcg gccccatca	33720
cgttgaagt tgatccaagt gctgagtttag atgtggtttgc gctcgtgtt ttacctgtct	33780
caccaatgcc cgtcaaaatg ccgcataatgc ccgcattttg aacgattttg ccgttctca	33840
taatggcagt accgcgaatg ccattgcggc catagttaaa tacggcgcgc tcgacgggaa	33900
tgcgtatgt ttcatcttca aaaatacagc tgaggcgagt ggtgcgcaca atgcggcgac	33960
tgcgtatgt ttgcggtaa gccgcgccta aaactgtaca gttgttttagt tttggggct	34020
tgcgtttagg cagtagcacct tggtaatgg tttggaaatac cattggtagc gatgcgcctt	34080
ggccattcac acccgcatgg gcatctgcgc cgccagtcac gacagctgtc acaaataaac	34140
ctgcaggaac gtagttctca ggggtgcgtt tttttggcag gttttggctt ggcgtttccc	34200
aatttagtgc aaaggaatgc atgcatttt gggttgcagg tggcgggata gcaccgcctt	34260
gatagggggc agcataccca ctttgtccgt caccacgtt cggcgtgg ggtacgttt	34320
gttcaccaa gtagcgtgtt ctttgctaa cgggataggc tgatgcata atgcctctg	34380
tgcggatgttcc atcggtgaca ttgttatgtc gcatgttgc tgcgttagc gcatataatct	34440
tatcatcaat tctttcgtt atggcgtttt ggaacgtctc taaattacgt tttaaatcaa	34500
ggcggttagc ttcatggctg cgctcaatgg cttgcgtgtc gtcagcattt ttatcaatcg	34560
ttttttgttgc ttatctaag gcaattttgtt gtgcggtagt cgcagactga ttgtctttct	34620
ctgtgaaatc ttatcaatc acggcacca aatcgattgg attgcccgtt gcggtttcag	34680
gagcacgtt agggcagac atataggcgt cgcccgctaa atagaggggca ccgagtagcc	34740
tggccccac ccctgcaatc acccaattgc gttttttatt ggcttcatatt aacgcactgc	34800
cttgctcaa gtcgggttcg gggttttttt tgaatgttgc taggagttgc ttcatggcat	34860
gtctcccccg cctgagacaa caaagaggtt gcgtgtttgc tgcgttgcata attcataggt	34920
tgataaggct gccgcgtcgcc tcgttggggc gtagactgc gccgttgtga gggtgactgg	34980
gagtgttgcg tgattggttt cgcgataat gatgcgtttt aatgttttac cagtgaaacac	35040
gaaagcagggt tctgttgtca gtcctgtgg tttggcatt tgctcattt aacgcgttacaa	35100
cataagggtca tctattgtat tttctgtgc tcttggctcg ttatgttgtt gttctgttc	35160
agcaggggca ctaactatgc acacacgtata accgtcaatc ggtttgcctt tgcgttgcgtt	35220
acggatcatt tgcgcggtaa actccgtcat catgggggggtaa taagggttt tcttgcgttca	35280

-continued

-continued

gattactgta ttacagtaaa cttactgtat tacagtaggg tggtaatac aaaatgaatt 37680
 attatgaaaa aaaattataa ccctcctact tcgtctaaa ctgagcttgt gacgtttagg 37740
 tgcccaaaga agctaagga attgatggac caagctgtaa aagatggtaa atatcaaaca 37800
 attacagctt taactgttga ggctgttaag gataaattgc aattcgaccc cgattcagat 37860
 gcatagttga tctgtttgc attggccttg caattttttt gatgagctat taaaatacga 37920
 agtagatttt gaagcaaaag tgcattacat agcacgtac tacatagtt tctatgtagt 37980
 gtatTTTca aaatttagagt gtgtggggg tggttcctg acgtgaagca ttgcacaaca 38040
 aagccgactt accgaatggg tcggctttt tgtgcctaaa attcaagaca ttgaaagatc 38100
 attgatctt gcgatcgatc atattatgt gtttggcaat gtttgaacta accattcaaa 38160
 cagccctaaa acattcggaa tcaacaataa tgcttaattt cttaaatggc tgctatttgc 38220
 agcacaagat ccatatcctc agatacatat accctctatg ccgccagatc tcgcgatcta 38280
 tgctgagcta cagtgggata gtgtctgtat caaaaactcta gtcaatcaga tcttcttat 38340
 tcatcattca gcagtgatga ttatcaagaa catcctctt tatctagtt tgctgatata 38400
 cctcatgttc atatggcaga actgccatg ggcaatgctt actctgtcg tgaagaagcg 38460
 attaggtaa aagttgctag aaaagctgag cgtgacaaga tcaatgccc aattaagagc 38520
 cttgctaaca tcgtcaagaa aaaagggtgtat gctgttggtt ttaaagcaaa agtagcatct 38580
 tatgctggat ttaaagctaa ggcactacct gatTTTGTGc aaacgttatt aacccatct 38640
 aaactattgc ctcaccgcga cgatTTTtaa gatcgtgggt atggattgtg tagagaaaa 38700
 actcaggcac gttctaaagt gatTTTGGC tatatgtct ctgcattttgt ttttaactgt 38760
 aatgtcaaca acggtcataat tgtcgttgct actagacatg gcccgaagaa cgttactcat 38820
 gatgaactac gtaaaagaagt ggccatgcgt cacgggtgtt acataccaga atcaacatgg 38880
 tactttacg ttaatcgatt agttcaatgc ggccacattt gaaGCCatc tttgagcatt 38940
 tatgaagatg gccacgttagc ctctttcat gctgaagcga gccataagta tttatcaaca 39000
 aagcttatgt caatgttggg tgctgaccga ttatctgtaa gaacagcagc tgctgaagcat 39060
 aacgttgc taaaactggc aggaaaatca tttaaagcaaa aaccccgctt cgcgatTTG 39120
 cgctatcgatc gcgatggttc tcttcgtcaa aatatgc当地 tcataaaaacc gtctcaagaa 39180
 ttgttggcac ttatcgctc tcactatgaa cgcgaagctt actacgtacc tcccaattga 39240
 cattctaagg ctTTTtaattt ctgcactat caatcgtac attactgtat ctTCTcatgt 39300
 cccaatgtcc aataagcccc tcataccctt aatTTTattt gatTTTGTCC atcaattttt 39360
 aatattactt agccaaattt gattggattt atcttattttt atgtatagcc atactttttt 39420
 ttaataaac ttaaagtaca aaaaaaaaaactg caaaatgtatgt ttttgcattt agtctcacct 39480
 ctTTTgattt ttTCTTTGta ttTTTGTCTT tagaaaaatga aagataaaaaa gatctatgc 39540
 aaaaatataa aaaaatgtatca ctgataaaatc agtgcgttctt acaagagatt aagagaagcg 39600
 ctgcaaggta atTTTGTCTT cgaccgttaa actacaaacg tgctgtact cagttgtaa 39660
 agaaaaaaaggc gtcttagcagg actagTTTGC aagctcaatg ccaaggatcataaaatagaat 39720
 ggttaaaacca ttcttattgtat aagtcattt gttaaagtaa tccctcaga tattcaacga 39780
 tccactttat agttaacgtt tagagaaaga tcaaaaacac cgcttaacacc tcctttgtc 39840
 atccgaccgt ttaacaacaa acggcgttgc caatcgttaa aaacccaaaaa acggcttgc 39900
 agggaaagaa atttcaaaaca ctgtgccatt aacaatcgta cggttttagt acggcttaat 39960
 ttgactaaaa tatcatttag acgtactatg tacggagttt tacataaaga taaaacacaa 40020

-continued

cagagtttgc tcgcagcaag taatgaggta ctaaaaaatg gaaatttga ataattttaa 40080
tgaatttagac aaatttattt tcttagattt ttattaagga ggttacaact taaaaaattt 40140
gatccatgtt ggaattcgat atgagtttc aaaatacggc aatgtcttgc gccgttcaa 40200
aaccagecgat gatagaactt aaaacctcgc tagagttaa ggaactgtttt gaacgagcgg 40260
cagcttatcaa tggcatcaat cttagggcat ttatcattaa tcaagcccga gagagagccc 40320
atgcaattat tgaatcagaa actacactac atctgaacca acatgtttttt actcaattt 40380
aaaccatttt agataatcct cgtaaagcta ctccggact aaaggacta ttttggagt 40440
aaaataatg agcacaagcg atggttattt tcagctgtt aacaagatgg aacgacaacc 40500
aagtttttca aatttttattt gtggtgatct gtttcttgat agttttgcgc ctaagaaattt 40560
ggcgaaagct gatgctaata atgattctcg ggtctatgtt gctgttagaca gagatatagg 40620
tgttggattt gcaacatgtt aagttttttt gctcagttat gacgagcattt caattttatc 40680
tggaaatat ccacgccagg ttctgtgtt aatgttagat caaattgcag ttgataaagc 40740
ttaccaaggc aaaggaattt gtaaacgcctt gatgcggaaa gtgtatggaa ctacgggttt 40800
agttatgag ctgtgtgcgc cgaagggtt agcggttattt gtcaccccta gagctaaaga 40860
tttttatata tcgtttagggt ttgtatccat tcctgtatca accaaacaag tgcaagacgt 40920
tgaatttagct ttaatgttta ttcatgttga aaccattttt gatgttatttttga agtaactaag 40980
agttatcaag aaacggattt ccaccactaa accccgcattt cgggggtttt atttgggtt 41040
ctgtgtccggaa accggccgaac gccaaaactt atgcctcactt gaaatttgcg cgggttccgac 41100
gtacacccccca taaaagcagt gctgttattt aattttttgc acgaccgtt aatattcaaac 41160
gcgcgtgtata gtcagtcgtt aaagaaaaag cgtcttagcag aagggttttga taagcttaat 41220
gcacaaatggt cgaagaaacg gagtttttttca aagcaaaag agttgtatcta 41280
tataaaatgg ttgaaccattt ttattgtttaat ttctacttgcg tatgttgcacca gtaagtaaga 41340
aagaaaggtt aaaaatttgcg aaggttattt agacctggca ttaaagagca ggaacttaat 41400
ccctgtcttc tatgggtttt tttgttgcgtt ttgttgcgtt aatacttcgc tttgttgcacc 41460
aatcagggtt cgttactaac cattgtatca acaatcccgat tgggttgcattt ggttacttca 41520
tttagacgcgc tttttcaattt tggttataactt gaaatttgcgtt cataagttttgc gtttacccat 41580
tccttaggtt tggcacaggc cagtgtatgtt cttggaccatc atcattggcc gtttttttgc 41640
ccgtcatttc tcatttcaactt accctcaactt tccgttatttgc gtttgcgttgc cacatttcaaca 41700
gtctgacccaa attcgggtt ttttacccat taaaatggca ttccatttttttttgc tataagtata 41760
caccttaaat attaaggat ttttacccat ttaatttttgc aaggaggattt tcatggat 41820
catgcagactt acagagactt ttcaagagct taaaatgggg gctgttgcctt atattaaagcg 41880
acgttacccaa cgggtttatgtt ctaatcccgat gaaagagtttgc cggaaattttca cgcgttgcgg 41940
agcctttacc tatttttttttttgc ttttgcgttgc aacacttgcgtt aatatttttttttgc ttttgcgttgc 42000
ttttgttca accggccatgtt aagatttgcgtt aatggcttgc aacatagaag aatgttataa 42060
attacgagat ttgttgcgttgc ttttgcgttgc aatatttttttttgc aatgttgcgttgc 42120
ccagaaaatg cagggttatttgc ttttgcgttgc ttttgcgttgc aatgttgcgttgc 42180
tgctgttgcgttgc ttttgcgttgc aatgttgcgttgc aatgttgcgttgc 42240
tatcgatatgttgc ttttgcgttgc aatgttgcgttgc aatgttgcgttgc 42300
aggttgcgttgc ttttgcgttgc aatgttgcgttgc aatgttgcgttgc 42360

-continued

tcttgaacaa gtcgtttcaa acgcattctt acctacaact atacctaatac tccgtatTTT 42420
 gcccggcatca caaagtgata gagctattga aggttggTTT catgaacaag tatttggca 42480
 aaagttaaag ttccttact ctctttgaa cacgatcatc aatgctgtc aagatgaatt 42540
 tgatatcatc attatcgata cccctccctc attagggtat gcaacttata atgcataTTT 42600
 tgccgctacc agttagtTTT tcccgTTGTC catcacagaa aacgacatg atgctactT 42660
 ttccTatTTT agttataatcc ctcaagtgTG ggctttattG gCGaatgCTA atcatcgTGG 42720
 ttatgatTTT atgaagatt taattacaaa tcatcgCGat agcgctacaa caaccgatCT 42780
 aatgaatagt ttatacgatc atTTTGCcC ttatATgtac tcaaAtgaat ttAAACatAG 42840
 tgaagctatt cgtcagtcat ctTcgTTGT ttctaccGTG tttgatATgt ctaagagtGA 42900
 ataccctaag agtaaAGcGA cgTTccaaAG tgcacAGcaA aattgttATg aagtaaccAG 42960
 ccaagtccta agagatattG tgaacgtcTG gCGtgaacAG gagcaAGcat aatggctAAA 43020
 aaacgtgggg taatgagccc tctaggtat gctgtggGT ccgaagaAGc acaaataAA 43080
 gCAGGCCAAAG ctaatattGA gtctttaaaa CGCCAAATTa caactgaaAT tgaaaaAGTA 43140
 agtgaagacG taacgttATC ttctcaAAAT ttatTTggGT ttGAatCTGT aggtAAAAGC 43200
 ttcttatggc aattagcttC tggtgctacc gctacattTA ccgaagcaAC attatcatAT 43260
 gaacaagttc gcgatagtac ctatgtGACT ttCGatGTT acggggcgtGA ccaggcatta 43320
 ttAAATGcAG attctctaca agatctcgat tcattAGCTT tccAGCAATT ttACCCAGCA 43380
 gtcgctAGAG aagtGAatGG taaaACTCGAT gtGCTAGATG gttctcgACG cAGAGCTGG 43440
 ttttactGc AAAATGGTGA agttGATA tttcgcatac tggtaactAA agatGatATT 43500
 tcactttcAG atgctAAAGC tctAGctAAA cagctccAAAG ctgAAAAGA acataACtA 43560
 cgtgaaATTG gCCAACAATG tttatctttG gaaaaAGcGA atcctaAGAT tacacAGGcT 43620
 gaagtagctG ctcaacttgg aatgagtcAG gctggTGTGA gtAAAGCTT AAAAGCCGCT 43680
 aaggTcGATG aacgTTTGGT gaagCTTTT cctgtggcTA atgacttGc acacactGac 43740
 tatgctttGc tgAGtAAAGt tatggAAAGtC tatGAatttG aagatGAatt actatCattC 43800
 atcaatggTT tgactAAACA agttGTCattt attcaggctG aatattCAAG ggaAGAACGt 43860
 aaatcAGctA tcacAAAGC gataAAAGC gaacttcAGA tcgctaAGGA tatgAAAAGt 43920
 aaAGCACAGG ttAGTGTtAC taatcttGcG acatttGata gtcaggcat ttacGcaAGA 43980
 aAGCgttAtTA aaggacGcAA ctCGcttAt gaatttggcT gactgtctt agatattCAA 44040
 ctgcAGTtAG atGTTGctAT tgCAGAtGTT ttGAAAAAAA taaAtttAaC acAAAAtCAG 44100
 atTTTAAAAA aatttAGTTT tatgtcactG caattAAattt atttctccGA ttGTTGgtACT 44160
 acataactAA acgccttAGt tacaccttAA atattGcTA tttataACAG tagataACTC 44220
 aggactatCT actgttAtCT atctatttGc tGtAtAtGAc ttCGGTTGcA ggtgtacata 44280
 tgAGcttAA acaatGcAA ttTTTcAGtG gtggatTTT ccaAAAtCCC aGcGACCCCT 44340
 caAAACCCACA taaAtACtTt aacttGttA atAAAtCCAC ttAtGACAAt atAtCtAtCT 44400
 gatAtttATA tattttACTA agggccttG gTgtttAtG AAAGAAGTT ttaAtGcTA 44460
 tcgagacGtG ttagttGTT ttTtGCGGtA tgccttAAGA tgcgtGctCT ttAAACGAGA 44520
 ttTtGtGAc cataAtGcAG tacctGattt agccccGgtTt atttGtGAGA agggtcAtGA 44580
 acattacGcG caacAGttGG aagtAtAtCT gAgccGAGcG gAACGGGtAA aAgGGAttGc 44640
 tgtaAccGgt ctttAcGcTA gtggcaAGAG tacGtttCtC aatacAtAc aggttcaCca 44700
 ccctGAAcTC aagtAcAtTA atAtCtCtt agctAAtttC tatGAtGAGG AtAttAAAC 44760

-continued

agttggtaaa aacacattat tagatgctca acaacggccg acagttgaac gcattgaacg 44820
 tgcagtatta aaggcgttat tatatacgaga aagtgcata gagagccgag gctctcgctt 44880
 tgcccggtca ccttaacta acccatctag accattcgca atcgctgtca cgtttactat 44940
 aacgatattt gggattgcat tgggtttgc gaacctttat ggcacaaaaa agatattaga 45000
 ttgggctcaa ataaaactcaa ccaattttac tgattcaacc caccctttat attggttgc 45060
 agcttcatg gtgcgagtgc caactttatt acttgctgat ttgggtcggt atgtagaca 45120
 gatacgaata agtaagataa atccgtttag tggtgatata gaggtacagg gtaaaagtca 45180
 tgactctgtg ttcaatctat accttgaaga tattttggct tacttcggcc atgctaaaat 45240
 tgatgtggtt atatttgaag atttagatcg ctggtagtgc catagaattt ttgaacggct 45300
 tagagaattt aataaaagtac tcaacgatag taatgttgc agtcggcccg tccgtttat 45360
 ctatgtctt tctgtatgtg tatttgaggg tcctgatcg actaagttt ttgatgcgat 45420
 tgttccaatc ottccgtgtgg ttgcgggggc taatgcctac ccacaattca aacaattact 45480
 ggccaaagct agtatactacta tcgcagacaa tacaagtcat tggatgatc tggcgaaac 45540
 tatcacgcta tatattcagg agatgaggct tcttaaaaggc atttgtggcg agttttatt 45600
 gtacagaaaaa gtgcttgaac tcaataaacac ccaaggcgc gaaacaaaat tattggcggtt 45660
 tatcgegtat aagaattttgt atatgtatga ttttgctta tgccaagaag gttaaggctc 45720
 gcttggtaaa caaatacaac aaaagaaaaa ttttataact ctagagcacc aaatgctga 45780
 aaagcaaata ttagagctc gtcaagaaga gaaagaggcc gaggcagatc atttggctga 45840
 tcaagggtgaa ttggctgagt tgatgttata tagatgtaat tcatctaaag tcactggagc 45900
 tgattcttct tatcatccgc ttctctctat tgatgttatt cccctatatg gtgttggaaa 45960
 tcctgtggag gtgggtgaaa gaatgtttaa cgccacctgc agcgagtcat cgtacggaa 46020
 cgtaatggaa ataacaagga atggggggtc tcatggtagc ctcagaaaat ggtcagctat 46080
 gcttggatcg gcaatgcctg attatcaac acgtctagcg cggttggaaa atcgaatat 46140
 agaagctcgta caagegcggc agaaaaaaat aaaacaacta caacaggaac aacaggctt 46200
 gagtttgcata agcctagcac aatgcctaaa gcccctctc aaccaggcga cacctacgcc 46260
 atcgaatgac aagccaatgc ttcacgcatt tctcatagca ggatttatcg atgaggatata 46320
 cggcttctac ttgagtgccc atgttgcagg gcatcttact aaacaagata tggattttt 46380
 aagagcattt aaggggctga cagtttgcata cccaaattat gaatctggca actataaaga 46440
 gctggtagca tttatcaatg gggaaagcatg tgcttctct gcccataata acatcgact 46500
 gatagagttt ctgtcagact cgccaaatttgc tcatgtatgc cttttttca ttgatgtttt 46560
 gaaaaatcaa tttaaagacc actcacaggg gcttgcgcgaa ttagctcaac ctatctggag 46620
 taaaatgtc tttaatctc ttataggca ttggcctgaa gtgttgcata cattgcaaaa 46680
 taatcaagtg ttaacaacaa gcgagactgc agtattactc gtttagatac tacttactat 46740
 tgaggaagta agagaagatg agtcattttt cattcagtca ataattgtatgc cctgtgcgt 46800
 tatgcttgcata atagtcgcag atatgtggcaatg tgatgtatgc gcttgcataatg taatgtc 46860
 agccgatattt aaaatcgaaa ctgtgaatatttttaccaac ttatgtgg tcattcgaga 46920
 ggccactgaaatc taaagctgaa ctcagtaaccc ttatcgccca caatatcagc 46980
 aataagagat aaagacatca tcactttcc ttgtgcataatg actaaattac caattacaa 47040
 caaaggctttt aaatcatttc tatattcaga tattcatgag tatgcaagat taatctattg 47100

-continued

tgggtgaatt tccgaaatgc cagcaacagt tatagtttcg ctatggaaacg atgagtttaa 47160
atcttatttg actcgagaag agaagcttc tctgatttag gggctagact ttctgattga 47220
ggatctcaat attatggacc ttggtaattt caagttgata gccttagccg aagagttcg 47280
aattgcgcct acttggctca acgttagctgt gttataaaat ttctataaca acgttctaaa 47340
tgataaaaaaa atagataaaaaa agcataacga aagcactgat gagttaaaag tcatggtgct 47400
gaattttta caagcacaga gtacaaaaga tgcatattgc tacaaagggtg agaagttaga 47460
acttgaatcg gtacgcgtat agtttatttc atttatcgaa tctagtgaga taaacgcgga 47520
tagtttgct gaatatatag ctgcggtaga gtatcaatat ggtcctgata atattatgg 47580
tttgcgttag gatcaatcct ccataattttt aaagagagga ttactttac catcatttg 47640
gctttatcg gcgctaagag ctgaccagca aaataatact gctttagctt tgatcacat 47700
gaacgaatca atattctta aacttaatat tatggaagtg gatggttctt gtttaccaga 47760
gttggagtt tcgtctgagg attttgttga attgatcgatc agtgattctc tttcgtagt 47820
agcaaaaacaa gcgttattaa aaactcatca agtaaatatt tataaagata ttgccccgaa 47880
agaatggta cactgggtt taccttattca aaagcaagaa aaaatatacg gactgactct 47940
gataaaaaat tcagttctga ctgagcaat tgctattgtat ttaggcttga ttcctgagca 48000
gttactattg gaagacgatg ttataaaagc attgcttgc atgaagtcaa ttcttcgga 48060
tgacaaagta cgtttattga tagggcaat ggctcatcg aaagcgaaaa tcttgccaaat 48120
gatcaattcg tgggactcga agccagatag tttctatacc actagtgtat gaatattgaa 48180
tggattgcta gataataata ccaattacgc ttctaggatgt tttttgggtt accatgaagt 48240
cgtatcttc tgctctgtca ggcgtggacg actacatatac aattacttta agggtagtt 48300
ttattnnca gccactcttc aacgatgact aggttattaa aaccgttac tcgtaatagt 48360
attgttagtt taaagtagga aaatatcgat gttatacaag aacaaggcatt gagtaaattt 48420
gegaatgccca tgaggatata tggcgaagct catatgaatt ttaatcgat ttagtttagtt 48480
gatgcccgaag aggctttaga taacttagat cgagctatgg aagccaagct tgaagcttc 48540
catagtctat atgatgtaac aaaggccctc tttgattact ttgaccatgc ggatacagca 48600
attctcattt tactgagaaa tgcagttcac caccgttacc atctgtttaaaaacttgg 48660
aatcaagaga tgggtttaaa tgaagggcac aaaaaatacc ttgggttga gttttgttt 48720
gcaagtcacg acattttgcac taatggcat gaaatgaagc atttttacaa gcttgaagat 48780
ttctatctgc gaatagaccc atcacttaggg tcacctttagt tcgaggatag aattagtgc 48840
aaaagtcgag aaaaattgtt aaatcagtc gaaaatgatc ttagtttgg tgccgttaaa 48900
aaatactcaa acagtgaag atacccgcata aaacatgtgt atataaaatatt aattcttatt 48960
tatattncaag ctactgtttaa ggtttcaaa gcacttaacg aaaagggtgt aaagtttgg 49020
ggttttgacg caaatgcata caaagaagct ttacaaatg aacttagctgt agattggc 49080
tcgtttaattt attcgacaaat acggatattt taattaaatg aacaactgtt taagagtgtat 49140
tatcaatgc aaggattttc actattctat ctatggattt cgggtggatt gcgttgc 49200
cgtcttaaca ggcagttacg tttcttataat taaatggtagc tattttatgaa atttttgaa 49260
gcaggagttc aatataaaaga tcttggatgtt tctgtccatg cagaccgaga tgataatcg 49320
gatgcaacctt attatctaaatg aaaaacccac agcatttcgtt ataaatgatc cgtgttgg 49380
atcaagttt attcgctgtt acataatgtt agatataataa cccctacagt tagtttttt 49440
cattcaaaatg ttggtagtta tgacaatatac caagaaaaga taaaaactgaa aggcgtatgt 49500

-continued

tggtttctaa acgaaggtaa aattgaaatgc cttacaatgc attttttagt tctgtttaag 49560
cgcttttagtt taactttatc atctaattgtt ttactcgaag gcaagtcata taccacaata 49620
taaaagctaa caaggcgta tgccatatctt ggagactatc gtgaataatgc agcaacgtct 49680
aagaattggg tactaccgtt cctctgattt aaacgatggg aaaatagatt ttggtatctt 49740
aatgcagat ttttctttag attcttatag agagctttagt cgcaactcga aggtgattct 49800
tgattttacag ccaatagcaa ttttgtcga agctgttcaa attaatttca acgaattaaa 49860
ttcttagttt aatattggga ttaactttatc gaagaatatgc aaacctggta aagggggtgt 49920
ttttcaatat ctttcattga tgagtgattt gtctgtaaaa attacaattt ttataaacctc 49980
agcgaataca tttcttgta actcggatgc caatcttaaa aagacatctg agcattttga 50040
atggaatgaa tatcgaataa aactacataa atcctcattt tcttacagat tcacttatga 50100
gttaagaaac tattcgcaac atcatagttt gccaatatac tgcgttataatgc tgaaccaaga 50160
taagacgaaa gataaaataa ctttgcttgtt aaacatgaaa agagatgagt tacttagttg 50220
tggttataaa tgggttaaaa ttaaaatgtt catccaaatgtt tgcgtacgagg tcttcgatct 50280
ttatcctcat ttaaataaat atatgaaaat tatagaggag ttattctca agtataataga 50340
cgtaaagagt gctaaactgc aagaaacaat catctatccc gcaaaactaa atagcacttt 50400
taagtccctt gaaaaaagtgtt cttctgttgtt ttgttgggtt gaagccgaaa atgacaaacc 50460
tgtacctaaa catcatgaaa tgataccatt tgaaaattttt gattggttat taagtatttg 50520
taataaggtt ggttcttattt cctgtatgtt acgttcttataag ggaaaattttt ccgttgcctt 50580
ttttcactta gttaaatgtt atcgttacgtt tttttttttt cattaactggg ttcgttaggt 50640
ttcttaggaag tattatgttat ctattagaca gatcaaagttt gaagctcggtt gacatcatc 50700
tcactcgctc gaatgaaaatgc aacagctcat taatttgcattt gattactaat tcaaattttt 50760
cacatgttat ttgttgcattt ggtgaatcga gctatattttt ccgttgcattt tatgggggtgc 50820
attcttagtaa tactcagcgtt ttactaatag atgagcttca atatgtaaaa gttatgcgag 50880
tggacgatcc cggttgcattt gaaaaagccat taaaatttttgc gcccgttacaa gttggaaactt 50940
cctattcaaa ggtttagtgctt gctaatgcgt tcgtttaatgtt tttttttttt ctagattcga 51000
aaaggcgatgtt ttgttcttagt ttgttgcattt aagcttttttgc gtcggcaggat gtcacaaactt 51060
ttgcaaaacat tgcgttgtt ttaccgcattt aaatcgttgc ttctatgtttt gttttttttt gttttttttt 51120
ttaaagggtttt tgcgttatcaat gctgaaccgc atgaaatttttgc tttttttttt gttttttttt 51180
ctatcaagaa tcaagctgaa attacaaaca acataactcaat atctgttgcattt gttttttttt 51240
gtaacaaat ccaatcactt tcagatattt ctaaaatgc tttttttttt gttttttttt 51300
atagcgaaat cactgtatattt tatgtgttccctt ctggataactt aaccatgtgg gttttttttt 51360
taaaagcaaa cccatggcgtt taaaatgc tttttttttt gttttttttt 51420
gtgaacaataa tgcgttgtt acacaagagt tgaaaatgc tgcgttgtt tttttttttt 51480
acaagattttt tttttttttt taaaatgc tttttttttt 51540
accaacaataa tttttttttt taaaatgc tttttttttt 51600
ccgaagcaat tttttttttt taaaatgc tttttttttt 51660
cgccctccactt cgacagtttgc tttttttttt 51720
caatgttgcattt tttttttttt 51780
atgttcttagt aatgggtttttt ggtttttttt 51840

-continued

ctagttcctt atatcgaaag tagaaagtca aagtcagagt caaaaagagc tattgatgt	51900
tttatgcgt agttggctga ttatgaaaag gcagccatct catacgtaaa aaattattat	51960
gaatcatact ggatacttg gaaggctgt aaaggtcaaa agttaatga tcgttttat	52020
gagcttaccc tagctcctcg tctagaattc ttatctttgg aaaccttgct tagcaagtct	52080
ttgcttaggc tgcgttagcga tc当地aaatgc gctgttaaag cactaaatgc tcttgagat	52140
aaaattaata aaaattcaga cctgattggg caagcgaaga gtgcggaaaga ttttgtgctc	52200
ctaattggaa aatttagagc aaatacagaa atgttagctt catttacta tttggtagt	52260
cgtatgcacatc atgaagctga tc当地ttgttct tatcttagatc taacaaatag tgaggtatgt	52320
aaaaaggtaa atggagttct tggcatttcg ttcaatggg aggtatctgc gaaagccat	52380
gtataacaaa ttgctgcacg ctgacacata ctgctacgct aattttgtg tttcgctgc	52440
gtcccatttt atacaaaagc actctccgca gcatgtgcag gtgagaaaga cggtgtcaaa	52500
aatgaaccag tgaatcactc tacaccagac tgatttcaccccgtatctt cttccaacg	52560
gggcctttttt tgccttagaa tgaatcaaattttagaa ttgattcatc gggttttagc	52620
tcatcaact tagcagccctt ggcaatggta gaacgactgc atttaagcag actctgcact	52680
tgtgaccagc ttcccttc atcgagtagt aaccgaatat tgcctcaatccatggta	52740
ggctgacgcc caaggattttt acctaagggtt ttggcctttt cgatgccttg tgcttgcgc	52800
tttcttttg tctcgtaatc atctctggcc atggccgcgg ctaaatcaat catgaactca	52860
gtcagacttcttgcgtatcatc ggtatgttct tcagtgatcattttaatccc atgtgtcatt	52920
ggctggtaa caacaacaat atgaatatacgcttttaccatattttt taggacccatc	52980
cattgctcaa aggggaggcg agacaatcta tcgacttttctatgatcaa aacatcacca	53040
gggtgagatg ccatgattaa ggcggccatgt tcaggccgtt caatttcgc ccctgatgc	53100
gtctcgtat aaaaactggc tatgcgaaca ttttactttt ctgaaaactgc aaccaactca	53160
tttctggctc gattggcatc ttgttcatct ttgtatgcgc gtaggttaggc ccgaataaac	53220
atttgcctc atttgggtttttttagatgt ggtttacgt gtagtgatgc gtaggtttttt	53280
aaatcaat atagcgaaacc aattggcaat gtatgtgtt tagtttgctt ggggtataacc	53340
ctaaatggtg ttgcgtctag ttgtatcctt gtgtatagat atactaatag cgacactt	53400
cggatatgag caacgtgaca ctactcaatgt tgatgcatacc tgatgcactt tgatgcctt	53460
cttgcggaaa cttacgatc atacttgccg caatgagaga gtaataact cttcaggag	53520
tatagaagct gtacttagat tatccccat cagcgataaa cacgtaaaca tgctggggca	53580
ttatcggtt acccttattt aacaagtact aagtggccaa ctcaggccgc agaagcaacc	53640
atcagagcg aacgaattga cttagcgtac gttttgttcc gttggacttc acactccat	53700
agcaaggctt aaccagtata aatatgctac gctacgccaa aaagtggccgc acttactta	53760
gggtgttatgt gcaatcaatc agaggtatcg tttttactt gaaatgaaaa ctttagttct	53820
tttttaatta ctgccttggc agagattgtt ggttgttattt taccataact ttgggtgcgt	53880
gaggggaaaat cgatgggtt gctgatccg gcccattatgttgcgtgtt gttgtgttgg	53940
ttgttattcc tacatccaaac agcagcaggc agagttatgttccgatccatgg tgggtttat	54000
atttttatgg ctatggatgt gctgtggta gtagatggaa tcaaaccac aacctggat	54060
ttgtgtggct cagggcgatc aattttgggtt atggtaatcaatccatggcc accacgcacg	54120
tagtacataa taagttgtt aattttgttc cggccacaaa aatcgtagcc agaactggac	54180
tggctacgct ggcggccatca attagcaag ctttacatc agatattttt cggccactaga	54240

-continued

tctgaaatcc taacaatttg tattccagag taatttacgc ttgttcgtt tcgattgtaa 54300
tcagttacca ttaaggcatta gtttccatag ttagaaagca aacttagaaa tattttcaaa 54360
gaggttcattt tttatcaaca acagagaatt gacatccaat acttgaattt agaatcttgt 54420
gttcagtcaa taaaatagat gtcactttgc gttattcacc aacgaaaaaa tagetcgatt 54480
attttatact tatcactgtat tttagccactg tagagttaa ggatgtcagt aatgaaacaa 54540
cgccatcatat taaaagctaa tattttcaaa gctctggggc atccaaacgcg atttaggtat 54600
gttagagcaac taacaaaggg tgaaaaatgt gtatgtaaat ttgttgaggg cggttatgtt 54660
gattttccaa caatttctaa gcacttatec gtttgcgcc atgctgaaat tgtggatgt 54720
aaaaaacgag gtaaggcagat tttttatcgt ctcaccatgc catgcttact gacatcgta 54780
cattgcattt atggtttattt agatcaacaa attcagaacg agatcgccatt aatggagtag 54840
tttcagactt atcatttggtaat taaatggcaaa aataacaaagtg taaccaagtg aggctaaatt 54900
atggatattttaa aatttctcggtacgggttgcg acaaagtgtc aaaaattggc agaggctact 54960
aatgcccgcg cgcaggact taatcttgcatttgc taccagctaa cttaaggatc tgatatttag 55020
aaaattatgg cctacaacgt aatgtcgaca ccggcgttgg tggtcgacga gcaataaag 55080
ctgtatggac gacttgcacg tggttatgaa ctcatgaccc tgcattcaagc ccatacagct 55140
tagtggcata aaggagtcgg catgaacgtt aaaaaatttgc tccgctaccc tgcgttact 55200
gtggtttgcgtt taggctttagg gttgggttggg tatcagcaat ttggcacgt aaccgctgt 55260
aaaactaaca atgcggatataat cgcgtccacgcgttgc acacgtgttgc tccgcatggaaaatttgc 55320
gtttatttttttcatgatggaaa tcagcgctgtt acaacgtgttgc tccgcatggaaaatttgc 55380
cagttactgttgcgttgc aatatttgc tccgcatggaaaatttgc 55440
ttgggttatgttgc tccgatttgc tccgcatggaaaatttgc 55500
cgttaccgttgc tccgatttgc tccgcatggaaaatttgc 55560
cgtgtgtggg agttggccaa caacgagctg gcattcagcc aatatttgc tccgcatggaaaatttgc 55620
aacgcataatgc ttaatcaaat tcaggcaat caaaccctcg cggcccccatttgc tccgcatggaaaatttgc 55680
ctccccccatgttgc tccgatttgc tccgcatggaaaatttgc 55740
ctctataatgc tccgatttgc tccgcatggaaaatttgc 55800
aatacaatgc tccgatttgc tccgcatggaaaatttgc 55860
gagtttgcgttgc tccgatttgc tccgcatggaaaatttgc 55920
cgttcgttgc tccgatttgc tccgcatggaaaatttgc 55980
gtttttgcgttgc tccgatttgc tccgcatggaaaatttgc 56040
aaagttctaaa atcaaccctc agggatattgttgc tccgatttgc tccgcatggaaaatttgc 56100
gtgtgttgc tccgatttgc tccgcatggaaaatttgc 56160
agagtcacaa tcgaggatattgttgc tccgatttgc tccgcatggaaaatttgc 56220
tttagttttcgttgc tccgatttgc tccgcatggaaaatttgc 56280
tattttgcgttgc tccgatttgc tccgcatggaaaatttgc 56340
tgccttatttgc tccgatttgc tccgcatggaaaatttgc 56400
taacaatgttgc tccgatttgc tccgcatggaaaatttgc 56460
agtacccatgttgc tccgatttgc tccgcatggaaaatttgc 56520
caatgttgc tccgatttgc tccgcatggaaaatttgc 56580

-continued

caactaccct ctgcagttag ccttttaaaa gtcacaaatc aaacgtttag ccagtatatt 56640
gtattgactg gaacaatttg accaacaaaa atggcgagtt tggcctctcc ggcagaagggt 56700
ccgattctta atcttgttgt gcgcgagggg gacactgttta atttaggaca agagattta 56760
cgcattggac gaactcatgc ggctaattct ttcaaacat ctgcccgtga agaagtaegt 56820
aaacaagttat taaatttaaa gcgtattgaa acttttagtta agcagcatac ttacactgag 56880
gaggcagttag acgaaagcaact acgtatcatta gaaaaagcga aggccggatt gagccaaagca 56940
aaacaagctc ttaatgatta catcggtacc tcgccttggg cagggattat atctaaagtc 57000
ttgggttcag atggtcattt cggttataccc cgtagttccct tggttgaat gtatgtatcca 57060
gatagttgg tattacgatt tagcgtcgca gaagcccaag cttagcgct taagaagggg 57120
cataaaattt aaggcacttt tgatgggctt gctggaaagg aatttgaatt agagatttt 57180
cgagcttatac ccgattttaga tagaaagttt cgtacgcgtc tttttgaggc ggcgtgcct 57240
gtaaatgagt ttattccagg tatgtttcg cggatcagt caataacaaca aaaacacgaa 57300
aatacccttg tcataccgtat tgatgcgtt caagtgcagaa gaaatgacaa aagtgtgttt 57360
gtgtgtactg ataataatagc cactcgacga ttatcgaaa ctggctttaga gcaggatgtat 57420
caaattgaag tggttatcggtt actgaaagcg ggcgagaaga tcgttctaac tggcattgaa 57480
cgggtgaaaa atggctctgc agtaagagta ttggagaaac ccgtaaaaaa caatagcgcg 57540
gggtgtgtga aatgtatcattt aaagccgcta ttctcgatcg aatttcaacg gctgtattgg 57600
cttcgggtttt agctatcccc gggcgctaa atctgaatat gtcacccgtc gatttcttac 57660
cttcgggttaa ataccctctc atcaagttt cgtatgttgc gcaagggtgca acacccgtt 57720
atatacgacca aaacttagcc gatecttattt agcgtgaact cgcctcgtt gatgttcttag 57780
attatctttc atctagcgcg atagaggccc tatacaattt ggtatgtcaat taccgttacg 57840
gtgtgtatgt tgatgttgcata tatacaagata cgttagccggc gtttacatgc tccaccaagg 57900
agctgcctgt tgatatttgc gctgccgtca tcatcaaaacg cgtatcgatca caattaccga 57960
tagtacaacgc tggtgttgcag tcggaaacata tggatgttgc tcaatcgatcg acttgggtt 58020
attcatgggtt aactcagcga ctgttatctg cgggtggcgtt agccgttattt gatgttgcgg 58080
ggggcttgc gctgttgcattt cgtatgttgc tggatgtatcaaaaactagaa gcccacggat 58140
tagatctcac cacgttgcataa cgggtccctcg ctgctgatggaa ttacaacccg gttggccggaa 58200
gggtgactgg ccaataccgtt gagaatatacg tgagggtgtat gggggagttt aggccgttag 58260
ctgttattca agatctcgat ttaagcccgata ataataatgg caacattgtt cgtgtcagg 58320
atgtggccga ggtttaaagat agccatgcg acattcgat gctgacccgc ttatgggtc 58380
atccggcggt taaatgttacatgtt gatgttgc gacgc tataactgtc acaacccgtt 58440
ataatgttgc gggccgttta gggggatgtt ccgtatctt tccatggat atcaattca 58500
cttttagtttgc aaaccaagcc gactatatca atgattctat ccgtgggtgtat cgtacactg 58560
ccttagaggc tatggcgctg gttgtgttgc tattattcgat attttaggtt aactcaaggc 58620
aagtgttataat tattgcgata gcttgcgtt ttgcattattt ggtgtatcc ttatgggtt 58680
acctcgccagg cttttccactc aatattttt cccttgggtt gtttgcgtt gcgatgggg 58740
tgttaccggc cacaatcaattt atagtggttgc aaaaatatttcc cgggtccgtt agtttacatg 58800
ataaggctaa cccacaatcc atctcccgagg aggctacactt ggaagttgggtt gggggcattt 58860
tggcagcgcac agtgcacccctt atagcgctgt ttgttgcgtt ttactcgat cctgggtt 58920
tcaccttactt atttaaagat ttgttgcgtt ttatccgtt gttgtatgtc attgcagggtc 58980

-continued

-continued

tcatcacaat	gccgacgcgg	aaataataac	cccaggtaat	attgatattt	ttctgcgcc	61380		
gtacatgcag	ccataacaag	gttgctaac	tgccgatggg	ggtgattttg	ggacctaata	61440		
cgc	tacgttcgcg	taaatcatgg	cctctttaat	cacgcccagg	gcattgctac	61500		
cgtctatcg	caggggc	cca	attaacactg	tgggcatgtt	gttcatgtg	gacgataaaa	61560	
aggcgacca	gaaggcctgt	cccata	gtgg	cccttgttct	gccaagctat	61620		
ttaatacgt	tgataggtaa	tcggta	ctgcgttgcg	aagaccataa	acaaccaaata	61680		
acatgcctaa	tgaaaagact	acgat	ttgcgcgc	gca	gactact	tttagtggat	61740	
ctatggcatg	gcctttttc	g	gcgacaacaa	acaggat	ggcacccac	gcagccacta	61800	
aactcacggg	cacaccgagt	ggttctaa	gccccaaagcc	caccaagagc	aagactaata	61860		
caccccaacc	ggtttaaag	gtgtt	gagg	cgcaatggc	ggcctttaggc	tcgcgttagtt	61920	
tttgc	taggttcgcg	ggaat	atctt	tgcgaaaaa	cagatgcagc	atgactaagg	61980	
tggcgcta	atgcggcaata	tcgac	gggc	ccat	actgc	acatattcg	ttaaagccta	62040
tgttgaagaa	gtcagccgaa	acgat	attca	ctaagtt	gacgattaac	ggcaactcg	62100	
cggtatcgc	aataaaaccc	gccc	catta	aaaggctaa	gtcgctccg	g	cgctaaagc	62160
c	agagccag	taacatcg	ca	attactatgg	gggtgaggat	cagcgc	ccatcat	62220
cgaacaatgc	ggcaacagta	g	gcgc	gagta	acactatgt	agtaaagagt	aaacggccac	62280
gac	cttgc	ccaa	acggg	cc	acatgc	cg	ggccat	62340
ataataaaact	gatgtatgc	acgg	caata	agg	ggcggt	cgcat	ttccag	62400
ataccacagg	aatatccc	ct	aat	gtgc	acca	caccc	cgtag	62460
agg	ttgc	act	cct	at	ggct	gcca	aattactaac	62520
gtacaaaat	tgccc	agct	aacc	cata	acc	at	ccat	62580
cgt	aaat	atgc	cata	agg	ata	tttgc	ccag	62640
cgc	gtta	ac	atgc	gggtt	tc	tttgc	ctgc	62700
cc	cttgc	tttct	ctat	ttgc	acc	at	ggc	62760
gagg	tcgtt	cc	ggc	tttgc	ata	ggc	cc	62820
ggc	ggc	aat	tttct	caat	tcg	ggc	caat	62880
gtg	acca	ata	catt	tttgc	act	gc	at	62940
tg	agttgtt	tttcc	cata	tc	acgg	gtt	agg	63000
ag	caataagg	tat	gc	ccgt	tttgc	act	tttgc	63060
cg	gataatgc	gcg	aaaaggc	tttgc	at	tttgc	at	63120
t	tttcttct	acaagg	ctt	gg	ata	ccgtt	tgct	63180
tg	ttcgcgg	tg	cg	at	tttgc	act	tttgc	63240
a	attgaccat	gc	agegtt	tttgc	at	ggc	ccgg	63300
t	ctaagccca	g	ctctg	ccaa	tctg	acc	atggc	63360
c	ccac	tttgc	ccat	gag	gat	ata	ggc	63420
c	ctatcc	atgc	gttgc	act	tcttgc	tgc	tacacc	63480
g	cggtt	aaat	ctat	gtt	tc	tttgc	tttgc	63540
ag	gttcatcg	att	ga	gtt	cgtt	tg	ggataaaa	63600
tt	tggtgagca	ccg	c	cttgc	acgtt	ggtag	caagggatc	63660
g	cgcc	tttgg	gca	ataac	gc	tttgc	acc	63720

-continued

-continued

gagggtaaaa acctgtttaa agcgggtaaa accgtcagta tcgagagctt taaagaacc 66120
 catacctacg gtttagtggaa gtgggtggaaac caagcttta aagactatac gccagaatgg 66180
 gcgagcaaaa ttactggat agaccctaaa accattattt ccattgctaa agatatgggg 66240
 gggcgccac ctgcggtgca agtatggact tcccgtggcg cagtgtgca agcccgccga 66300
 acatatactt cgatttcctt ccatgttta aatggctt atggctt cgatagtaaa 66360
 gggggtttat tcccaggtaa caaacgcct ctcctgaaag aataccaga ggcaaaagcc 66420
 tatatggatg agattgctgc taaaggtgtc aaaaaagaaa aaattgacca acgtggctgt 66480
 ttagcattcc ctgcatttc taaaggttaa cctggagggt ggggtgatcac gggtaacgctc 66540
 cctaattggca tgctcgctgc tgatcctt gaaataaaag tgattttagc ttatTTTaat 66600
 aacttcaact tttcaatcc tgaaggaaa cgttggatg aagccttaag caaagtgcgt 66660
 ttcatggccc atgtgaccac caacgtgtcc gaattcagtt ggttgcgtga tgtgttatta 66720
 cttctagcc accatatgtt tgaaaaatgg ggtgtgttgg attctatcgg taatgggttt 66780
 gcacagggtt caattcaaca accctcgatt aaacgtctat gggataccgg tatacgatgaa 66840
 tctgaaatcc cttacatgtt agctaaaaag ctggcagata aaggatttga tgcaccatgg 66900
 cgttatatac atgaacaaat tgcgtatcc gaaacaggta aacctgtgc ggtatggct 66960
 gagtttgcga agttgatggt cagattcttgc actgcgcgcg tggggaaaga agatgcgtct 67020
 aaatacggtg ataagctcaa ttcatggat gagttcgctgc aaaaagggtt gtggatagc 67080
 ttccttata agcttgatgc tcgctgggtt aagttcaaaa cagaaacccctc taagtttag 67140
 ttttacagta agaccttggaa aaaacgttg caagaacatg ctgataaaca caaagtgcac 67200
 atagatgagg taatgaaagc ttgtgactac caagctcggt gtcaatttgtt atttacccct 67260
 cattatgaag aaccctatcg atttgggtac gaagccgagt tcccgtgtt actcgtggat 67320
 caaaagtccatcgtc gtctaaacaa agaagggtcga actgcaataa gtccttggta ctacgagttt 67380
 aaagatgtcg atccctggta tggcgttggat gaagatgtgg cttaattccaa cccgatagac 67440
 ggcaagaaat ttggcctcaa agacgggtat gaaatccgca ttacaagccc tggggatcg 67500
 ctgacctgtttaa agggcgaagct atggggaaatgtt gtcctgttgc gcaactgtggc taagtgtttt 67560
 ggccaaggcatttttttgcata tggacgttac gccagttacca aattttggcat aaccccaaga 67620
 ggtggctcaa ataatgactt gattgcagac aggtacgtatc gtcataatgg tgcgtcgca 67680
 ttctatggtc atatccgtgt tcgtgtttagt aaagtgttgc gtaaccact atgagattag 67740
 gaatgggtat tgacctacaa aaatgtgttag ggtgtgggtt ctgttagctt gcctgtaaaa 67800
 cagagaacaa cacgaacgcgac ggtatttcatt ggtcgcatca tattgccacc actgagggg 67860
 tttttccatcgtttaa tgggggttgcata tgggggttgcata tgggggttgcata tgggggttgcata 67920
 cttgtgtaaa agtctgccccg acaggggtca tgcacaaaga taagcgccgc ttaacgtgc 67980
 aaaaacatgttgaatgttattt ggttgcata ggtgtatgaa tgggttgcata tgggggttgcata 68040
 ttatTTTaaa tgggggttgcata ccacatcgatc gttggcaggat gttggcaggat gttggcaggat 68100
 acggaaactgt atctccactg atgtgttgcata aacgcacggg ggctgtgcata tcacccaaatg 68160
 aaaaacccaga ggggggttgcata acttaccctg ttacacgcacc aagacgcact acagagaaat 68220
 gtactttctg tgcgttgcata cttgtataagg ggttgcata tgggttgcata tgggggttgcata 68280
 catctgttgcata ggggggttgcata atgtgttgcata ggtgtatgaa tgggttgcata tgggggttgcata 68340
 ttatTTTaaa tgggggttgcata cttgtataagg ggttgcata tgggttgcata tgggggttgcata 68400
 atatccgtcgttgcata tgggggttgcata aaaaacccaga acttaccctg ttacacgcacc aagacgcact acagagaaat 68460

-continued

-continued

ccaaaggta caacaaatag tccgataata atcgattca gtaatggact ggcaaacagg 70860
 aataccatca tagggccaaa accggcacga gccctgagta agccttcaa aaatggatg 70920
 gtagaacaag agcaaaacgg ggttaattgac ccaagtagtg cagcaatgac ataacctctg 70980
 ccgttgcgtg agcttagcat cgcttgtatt ttttggggg tgataaaatc ttgtaaaata 71040
 ctacttagt agcttatgac aagaatagt agggtaagct caacggcaag gaaagcgaac 71100
 atgcttgcgc tttcttcatacataaattc atttcaggac tcattttttt acctttcaa 71160
 cttaaagttt cgaatattct agaataatag tgtttggcg aagtgtttc aatatattat 71220
 ttctagaata atggaaatat ttggataacc tattgatgct taaacacatg acaccctgct 71280
 ttgattgtt tatgagtggc cgatttttagt gtgcgttagg taaggatatt tctggattta 71340
 tcgaattact taacgaaatg aggccgatatacataaattc attgctgcta aagtattaaa 71400
 agaactgggc catccaacgc gtctgcctt atttcgatta ctggtaagg gcggatatgc 71460
 tggcgttgcg gtgggtcagt tacaagacga gttcagatc cttggttcaa ctcttcaca 71520
 ccattattgtt gcccgtgtt cagcggagt gatttccaa cggcgcgagg gaagggtgct 71580
 gtactgcgtt cctgactatg aatttattaca aggatttagt catttttac aagatcaatg 71640
 ttgttagtgc cagtgaacaa ttttgtaaa tgaacagtaa attattacat gtggatattt 71700
 acatatatgta taaatcgat atagtgacat tcttttcatacataaattc tcccatgaaa 71760
 aaacgtgtac tttttctctg tggtggataat tctgcgcgtt cacaactcgc tgaggccta 71820
 ttaagacacc aagccaaaga gcagtttgat gtgttagt cgggtacgca gcctgagccc 71880
 atcgatgacgc gactcttgc gctttgcag aaaataatttgggtacgag tgagtacgt 71940
 tcgaagtctg tcagttagt tagtggcaaa tcctttgact ttgtgatcag tctctgtgaa 72000
 aaatcaacgc aagagtgtca gagttcccc ttggcagata agattattgc ttgggattat 72060
 cccgatccca aaatcgatc gggcactcgt gggttgaac aaacttttag agaactcaac 72120
 gaaagaataaa aaatgtttgt actcggttagt tctaaggatc ttaatgatcatacgc 72180
 attttttaaa tggttagccg atgaaactcg ctgcgtgtt tgatgttgc ttcagcacga 72240
 gggtagtgcgtt tgcgtctgtt aattgaccga agccttgcag gaaattcgc ccaagatctc 72300
 gggcattta gcccattac gtaagtgcgg attattagtc gatgcggc agggcaatg 72360
 gattttctat agtacgtatc atgatttacc cgaatggggg aagttagtgc tgagttagt 72420
 taccagccaa aatccgtat tccttgcataa aaacatcgatc aatctgtgca agatgggtgg 72480
 ccgcggcggc cgcgttagag cctgtgtca aagattgtt gtaacaata aataaaaaat 72540
 ttaatctaaa atatatgaaa attcatatatacataaattc taaatcgatc gaggttgcatacgc 72600
 tataaaaatgt ctcttcatacataaattc tgcgtgtcgc agcattctatc ctgtttttat 72660
 aggtcgagat ctgcgtcgatc agcaagcttt gactactatt gtcataatggc aatttgcag 72720
 tgccggtagt gaaacctgcgg gggtcgttca tcctcaact ctattgcac tagcacacag 72780
 aggctatgtt accgaaggcc tctgcagcaaa aagctggat atgatggccg atttcactcc 72840
 tgacttagtgc atcactgtttt gtgataatgc tgccggagaa acctgtccctt tggttttagg 72900
 tcaaaacactt aaattacattt ggggtttacc cggcccaaca tcaatcgacg ccccccataat 72960
 agatgagcaaa tttagtgcgtt ttagtgcgtt acgttgcgtt cgtataaagg cattaaatctc 73020
 gttgcgttgc tggcgaggta tagaagctca aaaagcatca ttacaatcaa ttgcgtgtca 73080
 atttccactt attcaagat aaatggtttgcgtt tattttatc aactaataa 73140
 aatagcaata tcctttgtt tcattgcgtt catgtataacc aaggaagttt gattatgtc 73200

-continued

caactatttt ccgatttagc gagctggcta acctttggag taatgggtt agatccaaat 73260
 actaagctcg ccgacgccc catatttttt attgaagata ccactaagat ttttgcgtc 73320
 ctgttgctga tgatttatgg catcgcttg gtgcgggcct cgctcaatgt cgagcgcgtt 73380
 cgggattact tggcggttaa aaatcgaaaa gtcgggtact ttatggatc gggttttggc 73440
 gcggttactc cattctgctc atgttcgagc attccggttt tttagggttc acctctgctg 73500
 ggatccccgt tggatcaact atggcgtttc tgattacttc gccgttaatt aatgaagtctg 73560
 cggttctgtt gcttgtgagt ctgttgggtt ggaagttac tgtatatac gtgtgggtcg 73620
 geatgtcagt gggatgttg gggggggcat ttttggacac gatccgcgtc gagcgttggc 73680
 tgcagtcctt tgccgccaata gcactcgagc aaggaaaggc acaagcaagt cacgataata 73740
 ggggggtat gacatcaaca tccatgacgt taacggaaacg gcatgaattt gcgaaaggcg 73800
 agaccctaga gatTTTggc cgagtgtgga aatgggtcat tattgggtt gggcttggcg 73860
 ccgcactcca tggatttgta cctgacgggtt ggatcgaagc ccacttaggc gatggtcaat 73920
 ggtggtctgtt tcctgoggcg gtattgattt gtattcctct gtattccaaat gcccacagggg 73980
 ttatccatcat catggagagc cttatcaact atggcgttgc cgtaggacaca acattggcat 74040
 ttgttatggc aacgggtgcc gcccaggtttctgagttcat tttgctcaag caggtgtatgc 74100
 aatggggttt actggccatc gttttggca ttttattgtt ttcatttacc ttaataggtt 74160
 ggatcttaa cgctataggt cccgttctgt gagaattata aaaatgctaa acatcaaagt 74220
 attaggcagt ggtatgtgca aatgcacaaa aaccgcttagt attattacgg ccategcacaa 74280
 cgaaaaagggc atcagtatttgcgttggtaaa gggaaaccaat ccagaagtca tcatgggtca 74340
 taagggtatgtt agtacacccg ctgtgggtat tcatgagaag ctatgtcattt gggggtccat 74400
 tccccataga gccatgattt aatcttgggtt agtgggttaa cacttaatg tcacatccat 74460
 ttgtatattt gccgttagag agtgggtgca ggttggattt tacccttgc ccagggacta 74520
 aatctgtccc tggacagag ggggtggca ttcttaaaggc ggcgggaaact gaggtcatca 74580
 taaccttaa gccaacttgc gaaattgcaaa catttggtc tgcattatgtt cccgtatattt 74640
 gccatgaagc ggggatccgt tgggtgcatt tacctataga agatgtgcg gcacccgcag 74700
 aggtattcga gctcggtttt gcaacgacaca aaggcagaact gtcggcatgt atgcaactc 74760
 aatccacaaat tgccattcat tgcgtgggtt gttccgggtc cacaggatata atggcggca 74820
 tcttgcgtttt actggcggtt ggcacccgttgg cagaagtgtt tacccaaatgtt caatccattc 74880
 gcccataatgc cttaaaccaat gtgcacatcaac gtggctatata cgaacagata acgctttaat 74940
 cttaaacaaa taatataaaa cacctaatat aaaactgagt gtgaaggatg caggtatgac 75000
 aatcaaaatc gggataatg gtttggccg tttggggactt ttagactgc gcgctgtt 75060
 gggctggaa gagggttggat ttgtgcagat taatgtccc gcccggatgtt cggcgacatt 75120
 agcccatatttgcgttggatccat tttttttttt gttttttttt tttttttttt 75180
 tggcgatgac atcattatcg atggcaagcg tattcgctgtt actcgcaataa aaaccatcg 75240
 gggaaacccatg tggatccatg tttttttttt gttttttttt tttttttttt 75300
 agcagtgtctg caagcttattt tagatcaagg cttttttttt gttttttttt tttttttttt 75360
 taaaagaagag ggcgtgttaa acgttgcattt gggaggtgtt catcaactctt atgacaaagc 75420
 tattcatccg attgtgactt gggcccttgc tactactaact tttttttttt tttttttttt 75480
 atgtatccac gaaaacccatg gcatagtgca tggatccatg acgactattt acgatattac 75540

-continued

caacactcaa actatTTAG atgcaccgca taaagatctt cgccgtcgcc gggcctgtgg 75600
 tttaaggcTTT atccctacga caaaggGTC agcgacggcc attaccata tttccctga 75660
 actcaaaggT aagcttaacg gccatgcggT gcccggTccca tttagcAACG cttcattaac 75720
 cgattgcgtg ttcgaggtga gtcgcaaaac gaccgaagct gaagtcaatc gcctgttaaa 75780
 agaaggcggca gacggaccgc taaaaggcat tttaggttat gaggaaacGCC cattagTCTC 75840
 ggTCgattat aaaaccgatc cgcgttgcag cattatcgat gcgctatcga ccattgattat 75900
 caatggcact caggtaaac tctacgctt gtagacaac gagtggggtt atgccaatcg 75960
 caccgtcgaa ctggcccgca tggcggTct gatggataag gcataaggctt tatggtaag 76020
 ctgacagggt tcttgctaa catatcaccc gagatccGCC agtattttgtt ggtcacaggc 76080
 aactattggg cattcacgct caccgatggc gcattacgta tgtagtgggt gctccatTTT 76140
 catggcttag gttatagccc gctgcaaatt gccatgttat tccctttcta taaaatctt 76200
 ggggtggtaa cgaacttagt cggcggctgg ctcggggcgc gtttaggctt aaataagacc 76260
 atgaatgttag gcctattttat gcaggatgtc gccttagca tgctgttgtt gcctagcggt 76320
 atgctcACGG ttgcttgggt gatggcggcgc caggccttgcg cgggtatcgc taaagatctc 76380
 aataagatga gcgctaaaag cagttatcaag ttgttgggtc ccaatgtgc tcagggttag 76440
 ctgtataagt gggttgccat gctcaCTggc tccaaaaatg cgctaaagggg cgccgggttc 76500
 ttcttggcgc ggcgcTTact gaccctgtt ggattccagc tggcggTgtt aggtatggcgc 76560
 attggcctat tactggtgtg gatTTTtagt ctgttaagt tgcaacgcga tttaggtaaa 76620
 gccaaaaaca aacctaagtt cacggaaatt ttctctaaga gtccggcggt aaatacgctt 76680
 tctgcccac gcatgtttt gtttggtgcg cgggatgtgtt ggTTTgtgtt ggctttaccc 76740
 gtttatttgg cctcagcctt tggtggat cattggatgc tggcgggtt tctcgactc 76800
 tgggtaatag gttatggcat agtgcAAggc tttgcacctc gttgacggg gacaaagtgc 76860
 gcgagccaa acaagggttcc cgtatggacgt agcgccttag gttggggcgc gatattgagc 76920
 atagtgcgg caggcattgc gctggcgata agttatgact tccatgcgc gaataactg 76980
 atttggggat tgatgtgtt tggcgcctt ttcgcgtatc actcttattt acacagctat 77040
 ctatcgtca gttatgcggA tgaagatggc gtatcgTTT atgtgggtt ttactacatg 77100
 gtaatgtca tgggacgcTT gatcgggact gtgttgcgtc gttgggtgtt tcaagtgtat 77160
 ggcattggcgg ctgtgtgtg gatacggcgc gcatTTatttgc gctcgcagc gcttatctca 77220
 attaagttcc caagacatag agcgtatTTT agacattaa agacattaa cgcggaaaaaa tgcacaaac 77280
 aatatgcgaa ttaacatata tgtcatattt catatgtatt ctatgtgt tacataaaga 77340
 cagttaaatg acacctatgt ctccaaatga ggatataat atgaaaaaaa cagcgttaat 77400
 gtcactgctt ggtttaggat tatttgcttgc tgTTGCTac gctagcgagt tcgatttacc 77460
 ggTTTGTc actgaagttag aagatggTcg tctatgggtt ttAAAGAGA attcggtcga 77520
 gttaactgag ttAAACAGC atggtgagcc agcaaaAGCA tttactgtca ttgggtgtgg 77580
 tcctaaggGGG atgacggTTT aaggcggccga tcaaataacc tttagatgaat atttagctaa 77640
 ggtAAAAGCG aactaatgtt cttatgttgc gttactcgat atttctaaat cccattttct 77700
 gatTTTtac gagttggAAC ttggTTcta actcgtaata tctaaaaatt aggtgaatta 77760
 ctctagccca gattgactga gtcagactta cctatcggtt ttacgtGCC aaaccctgc 77820
 tattttgtgc agccttctt ctaaccaatt ttggccccat tgcgtgtat gaccgatTTA 77880
 tcactcagtgc cacagaggat aattgactta catttagtca ttccaaaaat atggaagttaa 77940

-continued

agtcaaatta cgcctctata taagtgacca aattttagcgt accagtaaa gtcataatc 78000
 tgtggattt gaccagattt atgcaactaa agtggatttg tctgcaattt tttaccgtt 78060
 tccatcatcc attgccattt caagcgaaaat atcactcgcc accgaaccct gatTTTGT 78120
 ggataaacacc ccaaagagcg cggtgattat ccaccgaaaa aatcacgaac cggcagcagg 78180
 accgaaattt tagaatttt taaatttctt gctattatta ctgattatTT atacagtagt 78240
 ttggtgctat gaaactcata acatgccatg cgagtgcagg aatttgcggg tttccaagtc 78300
 ctgcagcggg ttacgttcaa ttacccctta gtcttgatca actgcttggg gagcatctt 78360
 getcgacctg gtttggcgt gctgcggggt gctcaatggg agggggtggg atctatgt 78420
 gegatataact ggtgattgac cgtgccgcta agegccgtaa tctctctatt gttgttgc 78480
 getataacgg agaatttacc gtaaaagctcc tcgatgaaaa ggcgggtta ctggttctt 78540
 tcgatcaaca gcaaaacatg acgtccgtt caatcgatga tgcaagatacc ttctcggtt 78600
 aagggggtgt catcaagtca atccgacttc acgagcattt aagcttaccc gaacaatatc 78660
 tgagtcaatc atgtacggac tgattgatgc taattcattt tatgtgtcat gcaatttgg 78720
 gtttagaccc gacttgcgcg aacttcccgc gattgtactc agcaacaacg acgggttgt 78780
 tggcagtc aatcgagcgg ccaagtccgt tggcgtgaaa aaatttgcctt cttacttgc 78840
 gctgcaacat ttgtgcagac aacataatgt gcaggtgttt tcatacttatt atgaatttgc 78900
 cgcagatctg tctgcaaaaa tggatcaagt gatcggcgtt ttgcaccccg agcaatacgt 78960
 gtactccatc gatgagtctt ttgtgtcgtt taaaggctgt gctgcaattt cagacttac 79020
 ggcacattgt gtcagactgc gccgaacagt ttggcgtgaa tgccgttgc cgggtgtgt 79080
 ggggggttgtt gagacattaa ccttggcggaa gttggctaat catgccccgaa aacaatttgc 79140
 tcaatataaa ggtgtttgcg ttattgataa tgatgcgcag cgtatttggaa ttttaaaatc 79200
 gatgccagtt gatgaagttt gggggattgg ccgcacactt acggtaaaac tgggattact 79260
 ggggtgtcat actgcctatg atttggcaca actggcgcgcg aaagtggcgcg gacaacattt 79320
 ttccatcgat gtagagcgaa ctgtgcgaga actcaatggg cagatctgtt aaacgtggg 79380
 tgcacccaaa gcagataagc agcaaatatt ttccacccgc agcctaggcg aacgaatatc 79440
 taccgggaa catttacacc aagegttagc aaagcatgc gcaattgcgg gtgcgaagg 79500
 tcgagaacaa ggctcattgt gtaaagcaat ggctgtgtt gcccactt cggccatgt 79560
 tccacagcct gtgtatTTTaaaggctgt tcagttact tgcgttgcgt atgacagtc 79620
 tgaactgtgc gccgcggta gcagcgagct atcaaacctt tatgcggcag gtgtgcgtt 79680
 ttaccgaatt ggtgtaggcc tgattgatct ttgcctaaa acgtcagtgc aatatgactt 79740
 atttaatgcg ccaaagagcg atcctaactt aatgaaaatc tttatgtatgc tgaataatcg 79800
 ctatggccgt gatgtgtgtt atattgtgc acagggggatg gatcagcatt ggactatgc 79860
 ccgtcagttt ttatcgccgc agtacaccac ccgttggtcc gatattccga tgattaaagc 79920
 gtaataggat tttatggtgc tttatccagg cttactcca cgtataacaac ctAAAACGTC 79980
 tgcataatTTTaaaggccatggaa aatgtatgc aatatgtatgc aatatgactt 80040
 gttaaaaaggc aacaagcgct aatggtaaa gagcaaggatgg acattattt aaagcgat 80100
 gaactgagtg atgctgacaa ggacaaaata gcaaggacg ttgctgtatgt tatcttgc 80160
 gcttgggatc tacaactcag tgatttgac ttggtaata ctgggttaaa acctggat 80220
 aacactaaac agtgttaaat atttctctatg gtcggcgttgc gtcacatcgt tctaataatca 80280

-continued

cggaaatgat ggacacagaa cgtcgatcta catctagagc tacatctata taaggataaa 80340
 acaatgagaa caaaaagaaga tttttcttt ggtaagacgc atcaaggaag caaaggata 80400
 acactccaat tgagegatat cgagtctttt gcggaaaaag tgtcagataa ttttttact 80460
 gcacaactga atagaatgct gcaggagcat gggggaaaggc tgactatatc agatgaaacc 80520
 tcgcctccca atttctggag cctcatcgat aaaattgcca tttagcaagt agggtcgt 80580
 gaaattatg ctgcgtatga cgtcaatgat agtgttaacg cgacgctggc ctgtgacatt 80640
 gtgctgttaa atggtgtgct ctcgattaag tcgcattggt ggcgtataa agaaattaga 80700
 gccccgtgaga ttgtttcttc gttattgggt cccttgacc tgaaagctct tcagaacaaa 80760
 acctatattc gctggatga tggcacaaca gagtcactgt tggagatgct tgattaccag 80820
 acagagctgaaa aaaaagggttt tctgtttgtt aaatatcccctt ctgttattaa tcgtgggtt 80880
 tcatacatca ttgatttggta gtgtgcgaca gataccggtc ggcgtggtat ttgcgtcagaa 80940
 ctggatggg atgtatatac agagcttcga acagaacgat gtgtcggaaa gtagcacatt 81000
 ggatactaaa gactgcaaaa aactaacggc tctaattgcgc tctggccgtt aatttggggc 81060
 gacaacttat actgtttaaag ttatgagttt tggggctgtt gatgctgtga gcgatagtga 81120
 aaaaggtagtc gttatagctc gtaacgataa gactggaatg cttgttatga acgaacaggt 81180
 ttagtgcggaaa aatgttagag acaatggtat tcctttgccg agtattcatg acataaattc 81240
 agagttttagtgc taatttagtgc ggaaaatcac gaatgttcaa ttcttaaga agaaaaccga 81300
 agatataat tcccaactata gatagggttt cagaagagat tcaagggttat cttcctgact 81360
 accaaggcccc gactataaat ttcgtttt tctccaaattc aaaatatac gggtcagtga 81420
 aaactgagtt gtggaaagag gataatctct tattggttca tgatattagg gcttccacta 81480
 caggggtaaa aaatggcact gctatggtca gttgggtggg ggctaactct tctgaaatga 81540
 tacaaccgt tcatgttatac tgggtgggt tgggttctg gtataaattg a 81591

<210> SEQ ID NO 2
 <211> LENGTH: 20
 <212> TYPE: DNA
 <213> ORGANISM: artificial
 <220> FEATURE:
 <223> OTHER INFORMATION: primer endA_L

<400> SEQUENCE: 2

gctgttgctt ccaatacgcac 20

<210> SEQ ID NO 3
 <211> LENGTH: 20
 <212> TYPE: DNA
 <213> ORGANISM: artificial
 <220> FEATURE:
 <223> OTHER INFORMATION: primer endA_R

<400> SEQUENCE: 3

ggcgctgcga cttactcatc 20

<210> SEQ ID NO 4
 <211> LENGTH: 24
 <212> TYPE: DNA
 <213> ORGANISM: artificial
 <220> FEATURE:
 <223> OTHER INFORMATION: primer She_Mph1103F

<400> SEQUENCE: 4

gaaaatcttgc agtagcgatc catc 24

-continued

```

<210> SEQ_ID NO 5
<211> LENGTH: 34
<212> TYPE: DNA
<213> ORGANISM: artificial
<220> FEATURE:
<223> OTHER INFORMATION: primer She_XmaJR

<400> SEQUENCE: 5
gttgttccta ggctgggcc atatcaacct ctag

```

34

What is claimed is:

1. An isolated plasmid comprising a nucleotides 63978 to 72599 of SEQ ID NO: 1.
2. An isolated plasmid pSheB, having the sequence shown in SEQ ID NO: 1.
3. A bacterial strain comprising a nucleotides 63978 to 72599 of SEQ ID NO: 1.
4. The bacterial strain according to claim 3, wherein the strain is a *Shewanella* sp. O23S strain deposited in the IAFB Collection of Industrial Microorganisms in Warsaw, under the deposit number KKP 2045p.
5. A composition comprising the isolated plasmid of claim 1.
6. A composition comprising the isolated plasmid of claim 2.
7. A composition comprising the bacterial strain of claim 3.
8. A method for selective removal of arsenic from mineral resources, raw materials industry waste or soil, wherein the step of dissimilatory arsenate reduction is carried out with a bacterial strain comprising nucleotides 63978 to 72599 of SEQ ID NO: 1, wherein the method comprises:
 - a) preparing the mineral resources, waste or soil and mixing with an appropriate culture medium enabling the cultivation of the strain,
 - b) adding an inoculum of this strain and culturing under conditions enabling growth and conduction of dissimilatory arsenate reduction,
 whereby arsenic is selectively removed from the mineral resources, waste or soil.
9. The method, according to claim 8, wherein the step of dissimilatory arsenate reduction is carried out under neutral or slightly alkaline conditions.
10. The method according to claim 8, wherein the mineral resources are copper deposits.
11. A method for selective arsenic removal from a variety of mineral resources, raw materials industry wastes or soils, wherein the removal of arsenic is carried out by dissimilatory arsenate reduction with a bacterial strain comprising the plasmid pSheB, having the sequence shown in SEQ ID NO: 1 wherein the method comprises:

- 15 a) preparing the mineral resources, wastes or soils and mixing with an appropriate culture medium enabling the cultivation of the strain,
- b) adding an inoculum of this strain and culturing under conditions enabling growth and conduction of dissimilatory arsenate reduction,
- whereby arsenic is selectively removed from the mineral resources, wastes or soils.
12. The method according to claim 11, wherein the bacterial strain is a *Shewanella* sp. O23S strain deposited as KKP2045p.
13. The method according to claim 11, wherein step b) is followed by step c) releasing the arsenic from the solution of step b).
14. The method according to claim 13, wherein the removal of the released arsenic is carried out by flotation.
15. The method according to claim 11, wherein step a) is carried out by: shredding and fractionation of the mineral resources, wastes and soils.
16. The method according to claim 15, wherein the mineral resources, wastes or soils of step a) have a fraction size of 125-250 µm.
17. The method according to claim 11, wherein the culture medium is R1-R2 medium (R1 salt: NaCl—1.17 g/l; KCl—0.3 g/l; NH₄Cl—0.15 g/l; MgCl₂×6H₂O—0.41 g/l; CaCl₂×2H₂O—0.05 g/l and R2 salt (KH₂PO₄—0.17 g/l; NaHCO₃—2.0 g/l; Na₂SO₄×10 H₂O—0.07 g/l mixed in a ratio 1:1), supplemented with sodium lactate, yeast extract and Tuovinen salts.
18. The method according to claim 11, wherein the culture medium does not contain NO₃⁻ and Fe³⁺.
19. The method according to claim 11, wherein in step b), the culture is carried out under anaerobic atmosphere conditions, and is carried out with flushing of the medium with a mixture of gases N₂:CO₂.
20. The method according to claim 19, wherein the mixture of gases N₂:CO₂ is in a ratio of 4:1.
21. The method according to claim 14, further comprising selective precipitation of arsenites with sulfides.

* * * * *